The Dissertation Committee for Karin Mia Keller Certifies that this is the approved version of the following dissertation:

Biopolymer Analysis by Electrospray Ionization and Tandem Mass Spectrometry

Committee:

Jennifer S. Brodbelt, Supervisor

James A. Holcombe

David A. Laude

David W. Hoffman

Sean M. Kerwin

Biopolymer Analysis by Electrospray Ionization and Tandem Mass Spectrometry

by

Karin Mia Keller, B. S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2004

Acknowledgements

Thanks to my advisor, Dr. Jennifer Brodbelt, for giving me the opportunity to do this research, and to the other members of the Brodbelt research group. I am also grateful to Dr. Gary van Berkel, Dr. Bob Hettich, Nathan VerBerkmoes, Dr. Josh Sharp, and the rest of the Organic and Biological Mass Spectrometry Group at Oak Ridge National Laboratory for a rewarding internship in the spring and summer of 2002.

Thanks to Leon Oehlers of Texas State University for assistance with drug/DNA duplex binding studies. Thanks also to Dr. Andy Ellington, who provided materials for aptamer/ligand binding studies, to Megan Breeden, who performed column binding assays in support of the same study, and other members of the Ellington group for useful insights on the behavior and significance of aptamers.

And finally, endless thanks to Brian Goolsby for encouragement, patience, good humor and unbelievable personal support.

Biopolymer Analysis by Electrospray Ionization and Tandem Mass Spectrometry

Publication No._____

Karin Mia Keller, PhD

The University of Texas at Austin, 2004

Supervisor: Jennifer S. Brodbelt

Electrospray ionization was used in conjunction with Fourier-transform ion cyclotron resonance (FTICR) mass spectrometry and quadrupole ion trap (QIT) mass spectrometry to study protein and oligonucleotide ions *in vacuo*. The results help to identify effective strategies for mass spectral analysis of these macromolecules and provide new insight on their gas-phase behavior.

Tandem mass spectrometry experiments were conducted to evaluate different ion activation methods for biopolymer sequencing. Multipole storage-assisted dissociation (MSAD) and sustained off-resonance irradiation collision-activated dissociation (SORI-CAD) were compared for protein analysis in FTICR instrumentation, and infrared multiphoton dissociation (IRMPD) and collisonal activated dissociation (CAD) were compared for oligonucleotide analysis in QIT instrumentation. In both studies, the differences in the observed fragmentation patterns were noted and the underlying reasons for these differences were identified. The relative utility of MSAD vs. SORI-CAD and IRMPD vs. CAD were assessed in terms of their ability to produce diagnostic information that could be used to identify the protein or oligonucleotide under study.

Tandem mass spectrometry was also employed to study the dissociation patterns of both DNA/metal and DNA/drug complexes. The preferred fragmentation pathways exhibited by these species were observed to vary with the initial charge state of the precursor. The effect of the oligonucleotide sequence, the identity of the metal ion, and the identity of the drug on these pathways was established and (where possible) interpreted in terms of the specific non-covalent bonding patterns present in the parent complexes.

Finally, electrospray ionization was evaluated as a tool for screening molecular recognition in nucleic acid aptamer/small molecule interactions. Gas-phase data for binding stoichiometry and relative binding affinity were compared with the known solution behavior for a series of well-characterized case studies. Any observed discrepancies were rationalized in terms of ligand structure and/or the nature of the intermolecular ligand/aptamer interactions.

Table of Contents

List of Tables

List of Figures

List of Illustrations

Chapter 1: Introduction

Life processes are driven at the molecular level by three major types of biopolymer: nucleic acids, proteins, and polysaccharides. Each is comprised of small monomeric units (nucleotides, amino acids, or sugars, respectively) and the number, sequence, and arrangement of these individual components dictate the form and function of the composite macromolecule *in vivo*. Detailed characterization of these biomolecules illuminates the mechanisms that propel replication, catalysis, and mutability, which are the hallmarks of living systems.[1]

A broad range of techniques is available for chemical analysis of biopolymers, based on physical phenomena as simple as the action of centrifugal force on a mass (sedimentation) to spectroscopic phenomena as complex as surface plasmon resonance (SPR). Mass spectrometry (MS), the analysis of ionic species in the gas phase, is a relatively recent addition to this arsenal. Biopolymers are large, polar molecules that would seem to be poorly suited to an analytical technique originally developed to study gases and small, volatile organic molecules. Because of recent technological advances in ionization and mass analysis, however, the observation and/or characterization of intact biopolymers by mass spectrometry is now routine, and the speed, sensitivity, and specificity[2,3] of mass spectrometric measurements can now be brought to bear on a number of fundamental questions in biochemistry and molecular biology.

This chapter profiles the MS instrumentation most commonly employed to analyze biopolymers in the gas phase, then introduces two key applications of MS in the study of proteins and nucleic acids: analysis of primary sequence and analysis of non-

covalently bound intermolecular complexes. It ends with an outline of the following chapters, which describe original research in the analysis of proteins and nucleic acids by MS.

1.1 INSTRUMENTATION FOR BIOPOLYMER ANALYSIS

For decades biological macromolecules could not be analyzed by MS because the known ionization methods were too harsh to generate intact molecular ions. By the 1970s and early 1980s, however, several new techniques capable of producing molecular ions for small proteins and nucleic acids had emerged, including field desorption[4,5] (FD), plasma desorption[6-8] (PD), and fast atom bombardment[9-11] (FAB). These methods made it possible to detect mass spectrometric signals for proteins with molecular weights up to approximately 5 kD, but they were rarely effective for larger analytes. However, the two critical breakthroughs in ionization methodology were made in the mid to late 1980s. In 1984 John Fenn interfaced electrospray ionization (ESI), a physical phenomenon reported years earlier,[12,13] to a quadrupole mass filter,[14] and in initial studies signals for proteins as large as conalbumin (76 kD) were generated with this new system.[15] In 1989, Karas and Hillenkamp used a technique later named matrix-assisted laser desorption ionization (MALDI) to ionize proteins as large as bovine serum albumin (67 kD).[16] A similar technique, reported at about the same time by Tanaka, produced gas phase ions for several proteins with the molecular weights ranging from 12 to 34 kD.[17] The use of ESI and MALDI spread rapidly, and both methods have been shown capable of producing gas-phase ions from macromolecules with molecular weights in excess of 200 kD.

Electrospray tends to produce multiply charged ions from analytes bearing more than one ionizable functional group, and even large proteins and nucleic acids therefore appear as an envelope of multiply charged ions at relatively low m/z regimes when

subjected to ESI. Inexpensive, rugged quadrupole mass filters and quadrupole ion traps, which normally operate up to m/z 2000, can therefore provide effective mass analysis of electrosprayed protein and nucleic acid ions in many applications. For additional mass range, ESI has been used in conjunction with time-of-flight (TOF) mass analysis, which can be effective up to m/z 350,000. Modern TOF instrumentation also offers improved mass resolution and accuracy relative to quadrupole systems, and to exploit this, a variety of hybrid TOF instrument platforms featuring a "front end" quadrupole mass filter,[18] quadrupole ion trap,[19,20] or linear ion trap[21,22] and a "back end" TOF mass analyzer have appeared. For the most exacting applications, ESI is often used in conjunction with Fourier transform ion cyclotron resonance (FTICR) mass analyzers, which offer the best resolution and mass accuracy available. To facilitate mixture analysis, ESI can be easily interfaced with techniques such as HPLC[23] or CE,[24,25] to provide "on-line" separation prior to mass analysis. Even highly complex samples can be processed and analyzed in this way; in one benchmark example, over 400 different proteins from E. coli cell lysates were identified by experimentally determined molecular weights after capillary isoelectric focusing.[26]

In contrast with ESI, MALDI tends to produce singly charged species, and protein and nucleic acid ions generated by MALDI can therefore have m/z values well in excess of 10,000. A mass analyzer with an inherently large range is needed to detect such species, so MALDI is most commonly used in conjunction with TOF instrumentation. MALDI is quite robust for mixture analysis because it usually produces just one peak per analyte, and signals for individual proteins have been generated from samples as complex as biological tissue using MALDI.[27] These signals reflect the analyte population at a small, specific location, so MALDI can be used in a rastering

mode to yield information about the spatial distribution and relative abundance of various proteins in normal vs. abnormal tissues.[27]

1.2 BIOPOLYMER SEQUENCE ANALYSIS BY MS

Mass spectrometry can offer a practical alternative to established wet chemistry protocols for determining the sequence of a nucleic acid or protein. Mass spectrometry was first employed in sequencing efforts in the early 1960s, when Klaus Biemann performed chemical digestion of peptides and derivatized the resulting amino acids to yield volatile species suitable for analysis by GC/MS.[28] Modern ionization technology produces intact multiresidue ions, so today gas-phase tandem mass spectrometry is usually employed to generate sequence information. Many fundamental studies of peptide and oligonucleotide fragmentation have been conducted using several ion dissociation techniques,[29,30] and standard fragment ion nomenclature systems for both peptide/protein[31] and oligonucleotide[32] dissociation have been devised (Figures 1.1 and 1.2). In general, both peptides and oligonucleotides tend to fragment between residues, and selectivity for particular interresidue bonds (the amide bond of peptides and the ribose 3'-O-phosphate bond for oligonucleotides) increases as the amount of energy deposited into the parent decreases. Small peptides and oligonucleotides often cleave at nearly every interresidue junction, so a fragment "ladder" can usually be constructed to retrieve the original sequence. As molecular weight increases, however, sequencing becomes more difficult: for peptides, sequence coverage tends to decrease and sequence information is therefore often incomplete, and for oligonucleotides, spectral complexity tends to increase and unambiguous data interpretation therefore becomes less likely. For these reasons, complete protein and oligonucleotide sequences longer than approximately 30-50 residues are not generally determined solely from mass spectrometric data. (It is

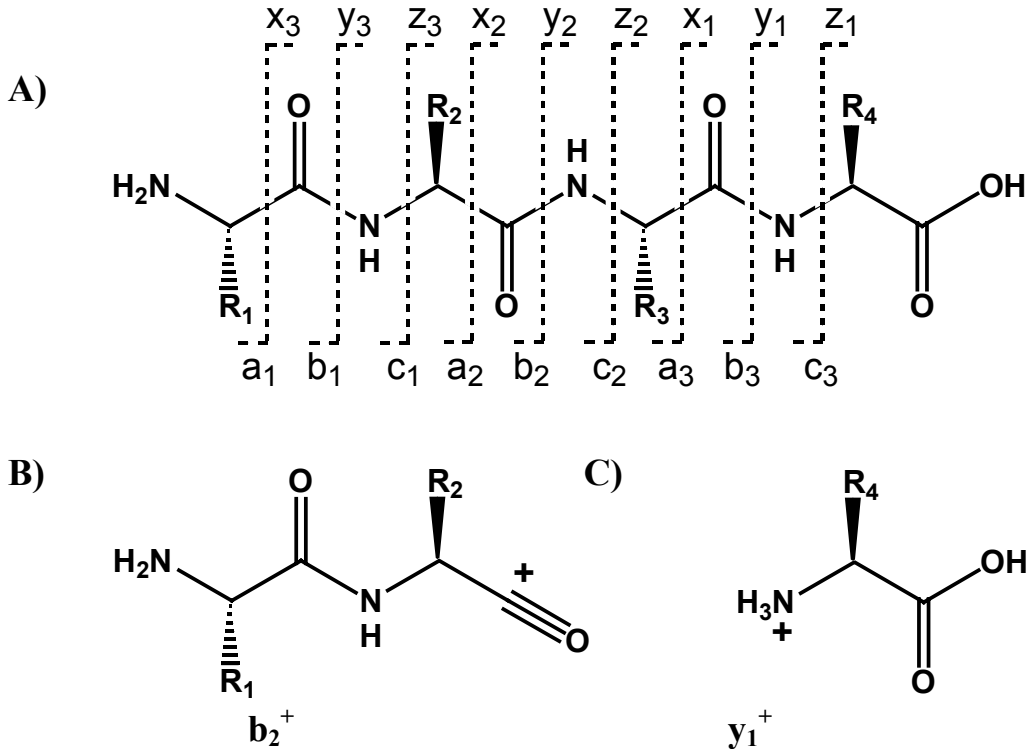


Figure 1.1: Fragmentation of peptides and proteins. A) Standard fragment ion nomenclature. Peptide fragments containing the N-terminus belong to the a/b/c series and fragments containing the C-terminus belong to the x/y/z series. Low-energy activation promotes the formation of b and y ions. B) Chemical structures of B) a representative b ion and C) a representative y ion.

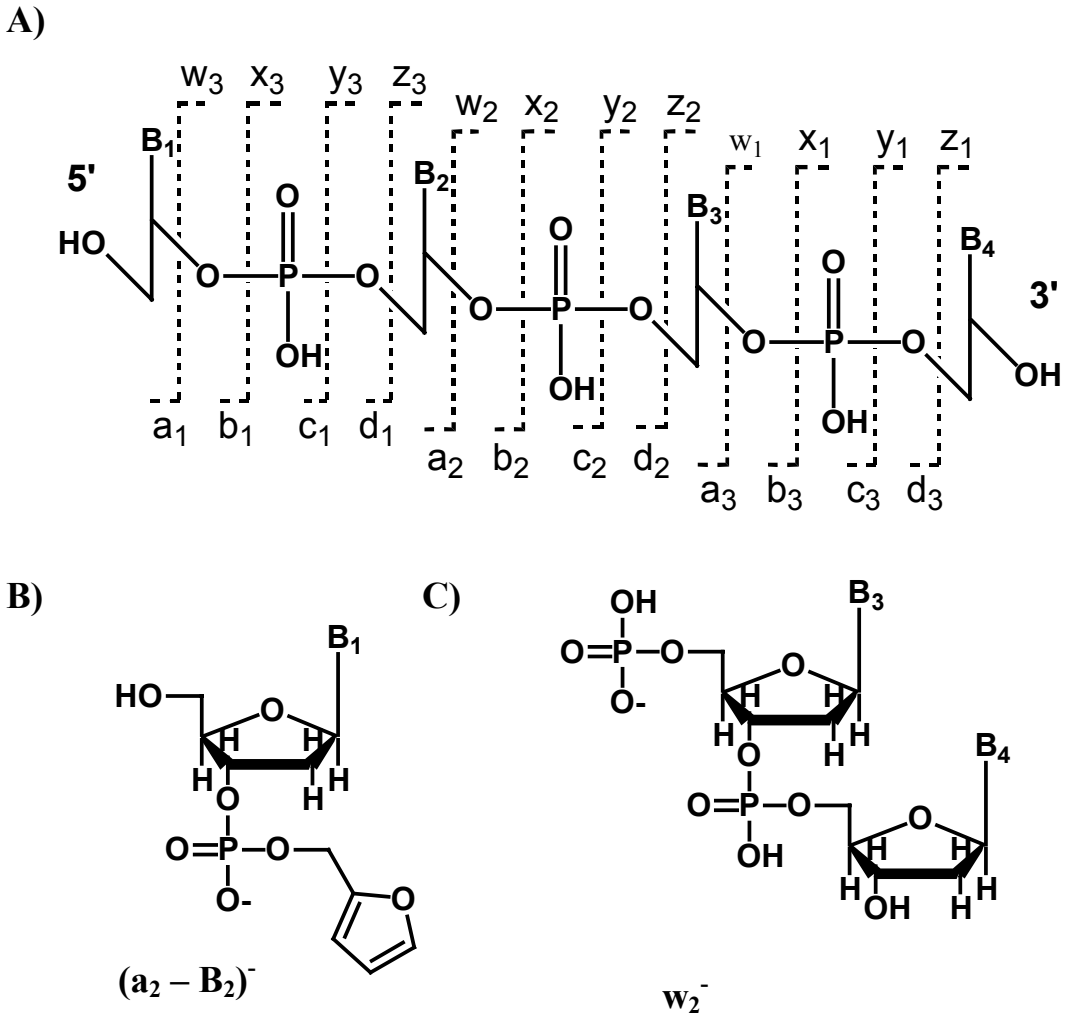


Figure 1.2: Fragmentation of oligonucleotides. A) Standard fragment ion nomenclature. Fragments containing the 5'-terminus belong to the a/b/c/d series and fragments containing the 3'-terminus belong to the w/x/y/z series. Low-energy activation promotes the formation of a_n-B_n and w ions. Chemical structures of A) a representative a_n-B_n ion and B) a representative w ion.

interesting to note, however, that a recently developed dissociation method, electron capture dissociation,[33,34] has been shown to provide nearly complete sequence coverage of a 29 kD protein.[35] This technique offers considerable promise for *de novo* sequencing but much development work remains to be done.)

Partial sequence information acquired by MS can, however, be used to confirm an expected sequence prepared synthetically or enzymatically, or to help identify and characterize sequence variants that may result from the omission, replacement, or modification of individual residues. In an important contemporary application, partial sequence data can also be used in conjunction with computerized sequence libraries to identify genomically encoded proteins from complex biological samples. Two distinct strategies have developed to address this task. In the "bottom up" approach, intact proteins are enzymatically digested to produce a mixture of smaller peptides, which dissociate far more easily.[36,37] Molecular weight and fragmentation data acquired for these peptides are cross-referenced with sequence libraries, and if the volume and quality of data is sufficient the parent protein will emerge as a "hit." The alternative, more recent "top down" approach is to record MW data for the intact protein, then dissociate it directly and use the resulting mass and fragmentation data to search sequence libraries.[38,39] This strategy requires high mass accuracy (generally provided by FTICR instrumentation) and different search algorithms,[40] but it is far more likely to detect post-translational modifications to the genomically encoded primary sequence.

In recent years, both the "bottom up" and "top down" approaches have been integrated with separation techniques, electrospray ionization, and automated search algorithms to provide protein profiling of highly complicated systems. To cite just one of dozens of recent examples, the bottom up approach provided identification of over 1900 proteins from *D. radiodurans*, 715 of which had previously listed as hypothetical.[41] The

top down approach is less mature, but significant progress in both front-end sample manipulation[42] and back-end data processing[40] has recently been reported. Integrated approaches aiming to exploit the strengths of both strategies are also beginning to emerge.[41,43]

1.3 MS FOR THE STUDY OF NON-COVALENT BIOMOLECULE INTERACTIONS

Biopolymers function *in vivo* by specific intermolecular associations with metal cations, small molecules, and other biopolymers, and these associations can occur on either a transient basis (as for enzyme-substrate interactions) or a more permanent basis (as for interactions between protein subunits or strands of DNA). Various aspects of these complexation processes can be studied by a number of solution techniques, including ultracentrifugation, microcalorimetry, surface plasmon resonance, and several different kinds of spectroscopy. In addition, X-ray crystallography can provide detailed information about structure and conformation adopted by proteins and nucleic acids when they interact with other biomolecules.

Over the last decade it has become clear that ESI-MS can also be used to study some non-covalent complexes formed in solution.[3,44,45] Early in the development of ESI, it was recognized that using gentle ionization and interface conditions (i.e. low temperatures and voltages) non-covalent complexes involving both proteins[46,47] and nucleic acids[48-50] can be transferred intact from solution into the gas phase. To generate and preserve these intermolecular interactions, however, the composition of electrospray samples must approximate physiological conditions. Buffers containing either ammonium acetate or ammonium bicarbonate are often used in place of standard biological buffers containing involatile metal salts, which are more likely to suppress ionization and can clog the sampling orifice of the mass spectrometer.[44] Organic co-

solvents can strongly enhance ionization in electrospray experiments, but because they tend to denature proteins and disrupt intermolecular bonding, experiments involving non-covalent protein complexes are usually conducted with strictly aqueous samples.[44] Interestingly, the intermolecular interactions of nucleic acids are more tolerant of (and may even be enhanced by) low proportions of some organic solvents,[51] so to improve instrument response studies involving nucleic acids are typically performed with samples containing 10-50% acetonitrile, methanol, or isopropanol.[51-57]

Using appropriate samples and instrument parameters, ESI produces intact gas-phase complexes at stoichiometries and relative populations that can directly reflect solution binding properties. ESI mass spectra have been used to observe the subunit composition of multimeric proteins,[3] distinguish between duplex, triplex, and quaduplex DNA,[58,59] and observe characteristic drug/receptor binding stoichiometries.[52,56,57,60] The relative peak intensities observed for complexes between a given receptor and various ligands can reflect relative binding affinities observed in solution,[52,61] and this type of analysis has been implemented in high-throughput drug discovery platforms.[62-64] In addition, absolute binding constants have been accurately determined by MS via titration experiments that relate the mass spectral peak intensity for an intact complex with ligand concentration.[60,65,66] In gas-phase dissociation experiments, the activation energy required to cleave non-covalent protein and nucleic acid complexes can correlate with solution binding affinity.[54,67-70]

However, gas-phase data for non-covalent complexes can at times present a misleading picture of solution binding processes. In some reports, the relative signal intensities[71] and dissociation energetics[72,73] observed in MS experiments with receptor/ligand complexes have provided relative binding trends that disagree with known solution behavior. Such results likely indicate that the relative influence of

electrostatic and hydrophobic interactions is different *in vacuo* than in solution, a notion which merits additional study in future research.

1.4 OVERVIEW OF CHAPTERS

Chapter 2 describes general experimental procedures employed in this research, including methods for sample preparation, ionization, and mass analysis. The subsequent chapters describe results obtained in gas-phase studies of the dissociation of intact proteins, nucleic acids, and certain non-covalent complexes of nucleic acids, as well as the association of nucleic acids with small molecule ligands.

In Chapter 3, two dissociation methods, sustained off-resonance collision-activated dissociation (SORI-CAD) and multipole storage-assisted dissociation (MSAD), are compared for the analysis of intact proteins in the FTICR. Because of non-resonant activation and mass discrimination effects, MSAD spectra often appeared quite distinct from SORI-CAD spectra. However, both techniques provided fragmentation information that allowed positive identification of the parent proteins in online database searches. Chapter 4 describes a comparison of infrared multiphoton dissociation (IRMPD) and collision-activated dissociation (CAD) for nucleic acid analysis in a quadrupole ion trap. The results showed that sequence information derived by these techniques was comparable in many respects. Diagnostic low mass ions not visible in CAD spectra could be observed in IRMPD spectra, however, and for this reason the use of IRMPD can sometimes facilitate identification of modified nucleobases.

Chapter 5 describes the dissociation patterns observed for a series of oligonucleotide/metal complexes in the quadrupole ion trap. Fragmentation data for low charge state precursors indicated that metal complexation may alter the gas-phase conformation of the oligonucleotide. Unusual fragment ions were observed for some

high charge state precursors, indicating that metal complexation may promote deprotonation of some nucleobase moieties. In Chapter 6, the influence of the initial charge state on fragmentation patterns for drug/DNA complexes is examined. Duplex strand separation predominated for higher charge states regardless of drug binding mode, while for lower charge states loss of ligand was more likely for intercalators and covalent cleavage was more likely for minor groove binders. The dissociation pathways for the lower charge state complexes are probably more reflective of specific drug-DNA interactions.

And finally, Chapter 7 describes non-covalent interactions of nucleic acid aptamers with their cognate small molecule ligands in the gas phase. The results indicate that the gas-phase data for aptamer/ligand complexes reflect the solution phase behavior in only some respects. Discrepancies were most likely to arise for complexes maintained in solution by significant π-stacking interactions.

1.5 REFERENCES

(1) Voet, D.; Voet, J. G. *Biochemistry*; 2nd ed.; John Wiley & Sons:, 1995.

(2) McLafferty, F. W. *Science* **1981**, *214*, 280-287.

(3) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(4) Beckey, H. D. *Research/Development* **1969**, *20*, 26-29.

(5) Beckey, H. D. *Principles of Field Ionization and Field Desorption Mass Spectrometry*; Pergamon: Oxford, 1977.

(6) Torgerson, D. F.; Skowronski, R. P.; Macfarlane, R. D. *Biochem. Biophys. Res. Commun.* **1974**, *60*, 616-621.

(7) Macfarlane, R. D.; Torgerson, D. F. *Science* **1976**, *191*, 920-925.

(8) Sundqvist, B.; Macfarlane, R. D. *Mass Spectrom. Rev.* **1985**, *4*, 421-460.

(9) Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tyler, A. N. *Nature* **1981**, *293*, 270-275.

(10) Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tyler, A. N. *J. Chem. Soc. Chem. Commun.* **1981**, 325-327.

(11) Barber, M.; Bordoli, R. S.; Elliott, G. J.; Sedgwick, R. D.; Tyler, A. N. *Anal. Chem.* **1982**, *54*, 645A-646A, 649A-650A, 653A, 655A, 657A.

(12) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240-2249.

(13) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287-2294.

(14) Yamashita, M.; Fenn, J. B. *J. Chem. Phys.* **1984**, *88*, 4451-4459.

(15) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.

(16) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

(17) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yohida, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.

(18) Morris, H. R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R. S.; Hoyes, J.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 889-896.

(19) Qian, M. G.; Lubman, D. M. *Anal. Chem.* **1995**, *67*, 234A-236A, 238A, 240A-232A.

(20) Qian, M. G.; Lubman, D. M. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1911-1920.

(21) Hager, J. W. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 512-526.

(22) Schwartz, J. C.; Senko, M. W.; Syka, J. E. P. *J. Am. Soc.Mass Spectrom.* **2002**, *13*, 659-669.

(23) Bruins, A. P.; Covey, T. R.; Henion, J. D. *Anal. Chem.* **1987**, *59*, 2642-2646.

(24) Olivares, J. A.; Nguyen, N. T.; Yonker, C. R.; Smith, R. D. *Anal. Chem.* **1987**, *59*, 1230-1232.

(25) Smith, R. D.; Olivares, J. A.; Nguyen, N. T.; Udseth, H. R. *Anal. Chem.* **1988**, *60*, 436-441.

(26) Yang, L.; Lee, C. S.; Hofstadler, S. A.; Pasa-Tolic, L.; Smith, R. D. *Anal. Chem.* **1998**, *70*, 3235-3241.

(27) Chaurand, P.; Schwartz, S. A.; Caprioli, R. M. *J. Proteome Res.* **2004**, *3*, 245-252.

(28) Biemann, K.; Vetter, W. *Biochem. Biophys. Res. Commun.***1960**, *3*, 578-584.

(29) Papayannopoulos, I. A. *Mass Spectrom. Rev.* **1995**, *14*, 49-73.

(30) Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *J. Mass Spectrom.* **1996**, *31*, 339-350.

(31) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.

(32) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.

(33) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

(34) Kruger, N. A.; Zubarev, R. A.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom.* **1999**, *185*, 787-793.

(35) Sze, S. K.; Ge, Y.; Oh, H.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U S A* **2002**, *99*, 1774-1779.

(36) Hunt, D. F.; Yates, J. R., III; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U S A* **1986**, *83*, 6233-6237.

(37) Yates, J. R. *J. Mass Spectrom.* **1998**, *33*, 1-19.

(38) Reid, G. E.; McLuckey, S. A. *J. Mass Spectrom.* **2002**, *37*, 663-675.

(39) Kelleher, N. L. *Anal. Chem.* **2004**, *76*, 196A-203A.

(40) Meng, F.; Cargile, B. J.; Miller, L. M.; Forbes, A. J.; Johnson, J. R.; Kelleher, N. L. *Nat. Biotechnol.* **2001**, *19*, 952-957.

(41) Lipton, M. S.; Pasa-Tolic, L.; Anderson, G. A.; Anderson, D. J.; Auberry, D. L.; Battista, J. R.; Daly, M. J.; Fredrickson, J.; Hixson, K. K.; Kostandarithes, H.; Masselon, C.; Markillie, L. M.; Moore, R. J.; Romine, M. F.; Shen, Y.; Stritmatter, E.; Tolic, N.; Udseth, H. R.; Venkateswaran, A.; Wong, K.-K.; Zhao, R.; Smith, R. D. *Proc. Natl. Acad. Sci. U S A* **2002**, *99*, 11049-11054.

(42) Meng, F.; Cargile, B. J.; Patrie, S. M.; Johnson, J. R.; McLoughlin, S. M.; Kelleher, N. L. *Anal. Chem.* **2002**, *74*, 2923-2929.

(43) VerBerkmoes, N. C.; Bundy, J. L.; Hauser, L.; Asano, K. G.; Razumovskaya, J.; Larimer, F.; Hettich, R. L.; Stephenson, J. L., Jr. *J. Proteome Res.* **2002**, *1*, 239-252.

(44) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.

(45) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(46) Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1991**, *113*, 8534-8535.

(47) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294-6296.

(48) Ganem, B.; Li, Y. T.; Henion, J. D. *Tetrahedron Lett.* **1993**, *34*, 1445-1448.

(49) Light-Wahl, K. J.; Springer, D. L.; Winger, B. E.; Edmonds, C. G.; Camp, D. G., II; Thrall, B. D.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, *115*, 803-804.

(50) Doktycz, M. J.; Habibi-Goudarzi, S.; McLuckey, S. A. *Anal. Chem.* **1994**, *66*, 3416-3422.

(51) Sannes-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264-271.

(52) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(53) Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464-467.

(54) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(55) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(56) Reyzer, M. L.; Brodbelt, J. S.; Kerwin, S. M.; Kumar, D. *Nucleic Acids Res.* **2001**, *29*, E103-103.

(57) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(58) Goodlett, D. R.; Camp, D. G., II; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, *22*, 181-183.

(59) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.

(60) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.

(61) Loo, J. A.; Hu, P.; McConnell, P.; Mueller, W. T.; Sawyer, T. K.; Thanabal, V. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 234-243.

(62) Hofstadler, S. A.; Sannes-Lowery, K. A.; Crooke, S. T.; Ecker, D. J.; Sasmor, H.; Manalili, S.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 3436-3440.

(63) Sannes-Lowery, K. A.; Drader, J. J.; Griffey, R. H.; Hofstadler, S. A. *Proc. SPIE-Int. Soc. Opt. Eng.* **2001**, *4264*, 27-36.

(64) Cummins, L. L.; Chen, S.; Blyn, L. B.; Sannes-Lowery, K. A.; Drader, J. J.; Griffey, R. H.; Hofstadler, S. A. *J. Natural Prod.* **2003**, *66*, 1186-1190.

(65) Lim, H.-K.; Hsieh, Y. L.; Ganem, B.; Henion, J. *J. Mass Spectrom.* **1995**, *30*, 708-714.

(66) Greig, M. J.; Gaus, H.; Cummins, L. L.; Sasmor, H.; Griffey, R. H. *J. Am. Chem. Soc.* **1995**, *117*, 10765-10766.

(67) Kraunsoe, J. A. E.; Aplin, R. T.; Green, B.; Lowe, G. *FEBS Lett.* **1996**, *396*, 108-112.

(68) Rogniaux, H.; Van Dorsselaer, A.; Barth, P.; Biellmann, J. F.; Barbanton, J.; Van Zandt, M.; Chevrier, B.; Howard, E.; Mitschler, A.; Potier, N.; Urzhumtseva, L.; Moras, D.; Podjarny, A. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 635-647.

(69) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, *120*, 9605-9613.

(70) Gabelica, V.; De Pauw, E. *Int. J. Mass Spectrom.* **2002**, *219*, 151-159.

(71) Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. *J. Am. Chem. Soc.* **1996**, *118*, 8646-8653.

(72) Li, Y. T.; Hsieh, Y. L.; Henion, J. D.; Ocain, T. D.; Schiehser, G. A.; Ganem, B. *J. Am. Chem. Soc.* **1994**, *116*, 7487-7493.

(73) Wu, Q.; Gao, J.; Joseph-McCarthy, D.; Sigal, G. B.; Bruce, J. E.; Whitesides, G. M.; Smith, R. D. *J. Am. Chem. Soc.* **1997**, *119*, 1157-1158.

Chapter 2: Experimental Overview

The experiments described herein were performed using electrospray ionization and commercially available FTICR or QIT instrumentation. FTICR instruments are known for their high resolution and high mass accuracy capabilities, whereas QIT mass spectrometers offer a high degree of ruggedness and sensitivity. An IonSpec HiResESI FTICR, located at Oak Ridge National Laboratory, was equipped with a 9.4T magnet and allowed the comparison of two ion fragmentation methods, sustained off resonance irradiation collision-activated dissociation (SORI-CAD)[1] and multipole storage-assisted dissociation (MSAD)[2] for analysis of intact proteins. A Hitachi 3DQ QIT mass spectrometer was modified for infrared multiphoton dissociation (IRMPD)[3] and thus allowed an assessment of IRMPD for fragmentation of oligonucleotide anions in the quadrupole ion trap. A ThermoQuest LCQ Duo QIT system was used to study dissociation of oligonucleotide/metal complexes, drug/duplex DNA interactions, and small molecule/nucleic acid aptamer interactions. Brief descriptions of the major techniques employed in these studies are provided below.

2.1 ELECTROSPRAY IONIZATION (ESI)

Electrospray ionization[4] generates gas-phase ions directly from liquid samples, and has been widely implemented in mass spectrometric investigations of biomolecules as well as an enormous variety of other analytes.[5] In a typical electrospray experiment (refer to Figure 2.1), the sample solution is infused through a capillary held at high voltage (1-5 kV). Charged droplets form as the sample exits the capillary, and these

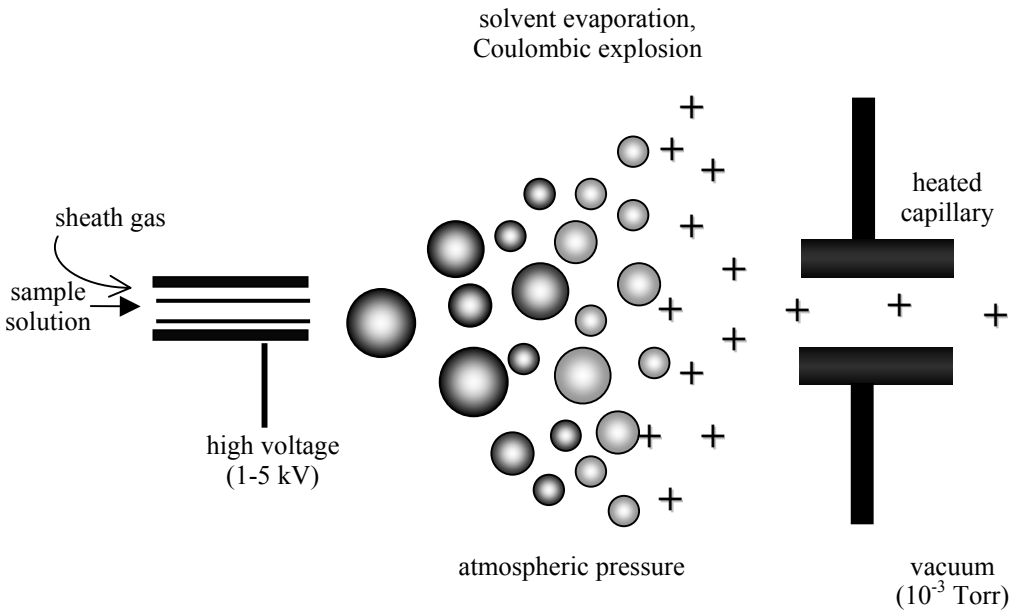


Figure 2.1: Schematic diagram of electrospray ionization.

droplets rapidly produce ions that enter the vacuum region of the mass spectrometer through a small sampling orifice. Heated elements in the front end of the mass spectrometer assist the desolvation process (a heated capillary is shown in Figure 2.1, but heated lens assemblies can also be implemented). Ions are then transferred through an ion optics region to the mass analyzer.

In these studies, electrospray samples contained a protein or oligonucleotide at concentrations of 2 x 10^{-6} to 2 x10^{-5} M in a mixed aqueous/organic solvent (10-50% acetonitrile, methanol, or isopropanol). Samples used to study non-covalent gas-phase complexes of oligonucleotides also contained secondary analytes, such as metal ions and/or "guest" molecules, in modest proportion. Various additives, such as acetic acid, the combination of piperidine and imidazole, or ammonium acetate were also employed as needed to promote ionization. The resulting samples were introduced to an electrospray source by continuous infusion at 1-5 μl/min. ESI needle voltages were typically +3.5 kV for detection of protein ions in positive mode and –2 to –4 kV for detection of oligonucleotide ions or complexes in negative mode.

In FTICR experiments using the IonSpec instrument (Figure 2.2), ions were accumulated in the optics region and periodically transferred to the ICR cell (typically 2 x 10^{-10} Torr). In QIT experiments ions were accumulated in the mass analyzer itself. Typical operating pressures were 3.5 x 10^{-5} Torr for the Hitachi instrument and 1 x 10^{-5} Torr for the ThermoQuest instrument (Figure 2.3). Both QIT systems employed a buffer gas (He) to facilitate trapping and resonant ejection.

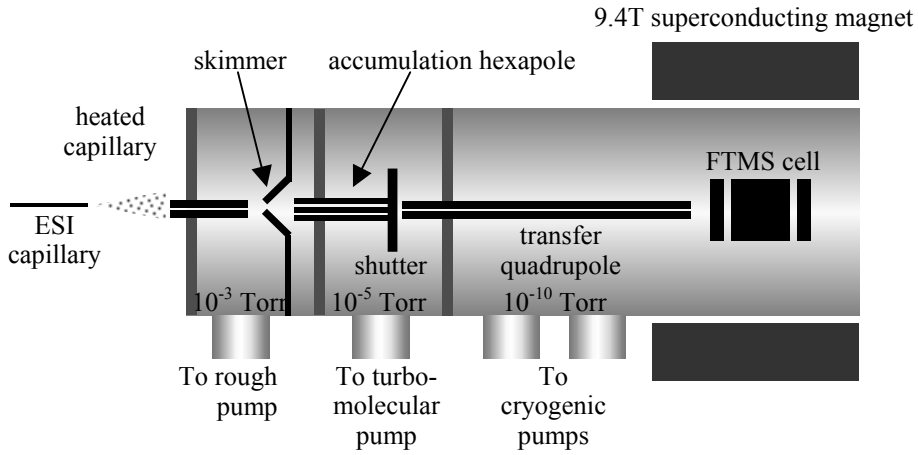


Figure 2.2: Schematic representation of the IonSpec HiResESI FTICR mass spectrometer.

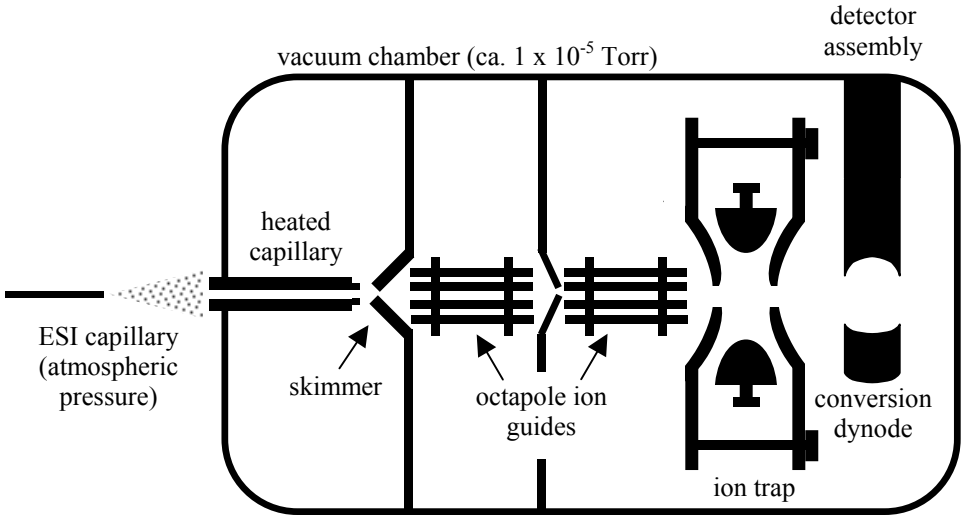


Figure 2.3: Schematic diagram of the ThermoQuest LCQ Duo QIT mass spectrometer.

2.2 ION DISSOCIATION AND MASS ANALYSIS

2.2.1 Collision-Activated Dissociation (CAD)

Collision-activated dissociation is a process by which energetic collisions with neutral molecules cause ion fragmentation, and is the most widely used method for structural characterization of gas-phase ions. In this research CAD experiments were conducted on both FTICR and QIT instrumentation.

A common variant of CAD, sustained off-resonance irradiation (SORI-) CAD was performed in the FTICR mass spectrometer. In this technique, non-resonant waveforms excite trapped ions to encourage collisions with a neutral target gas. In the studies described herein, precursor ions trapped in the cell were isolated with stored waveform inverse Fourier transform (SWIFT) pulses, and excitation waveforms were then applied at a frequency 1 kHz lower than the ion cyclotron frequency of the isolated precursor. The collision gas (N_2) was admitted to a transient pressure of 5×10^{-6} Torr in synchronization with the excitation waveforms. After a 10s pump down delay, mass analysis was achieved by broadband excitation/detection.[6]

Standard CAD was performed with resonant excitation waveforms in the QIT mass spectrometer. Precursor ion isolation was achieved by applying either broadband axial waveforms or by DC/RF voltages. Ion activation was then achieved by applying resonant frequencies across the end-cap electrodes for 30-50 ms to promote energetic collisions with the background gas (He). Throughout the experiment the voltage on the ring electrode was maintained to provide a q value of about 0.3. Under these conditions the lower mass range is truncated at an m/z value that is about 1/3 of the precursor ion. After the dissociation period, fragment ions were detected by mass-selective axial instability ejection with axial modulation.[7]

2.2.2 Multipole Storage-Assisted Dissociation (MSAD)

In MSAD, space charge-mediated dissociation occurs in the linear RF-only multipoles commonly used to interface electrospray ionization with FTICR cells. This method is a form of CAD[2,8,9] that provides efficient but non-resonant (i.e. not mass selective) activation. In MSAD experiments, electrosprayed ions were allowed to accumulate in the hexapole region of the IonSpec FTICR mass spectrometer. To encourage dissociation a DC offset was applied to the hexapole during accumulation. An electrostatic lens was then used in conjunction with a mechanical shutter to transfer the resulting ion population through the high vacuum region and into the cell, and mass analysis was performed by broadband excitation/detection.

2.2.3 Infrared Multiphoton Dissociation (IRMPD)

IRMPD is an alternative to CAD in which irradiation of trapped ions with infrared photons leads to vibrational excitation and ion dissociation. The commercially available Hitachi 3DQ system was modified in house to allow IRMPD experiments (Figure 2.4). The vacuum chamber was customized to position an IR-transparent ZnSe window (II-VI Inc.) directly over the ion trap, and a 6-mm hole was drilled through one side of the ring electrode. The beam from a continuous CO_2 laser (Synrad, Inc.) was directed through the ZnSe optic and into the center of the trap, and TTL pulses supplied by the 3DQ software were used to power the laser on and off at appropriate intervals. In addition, external DC inputs were applied to allow DC/RF isolation,[7] which provides greater ejection efficiency for high m/z ions on this instrument.

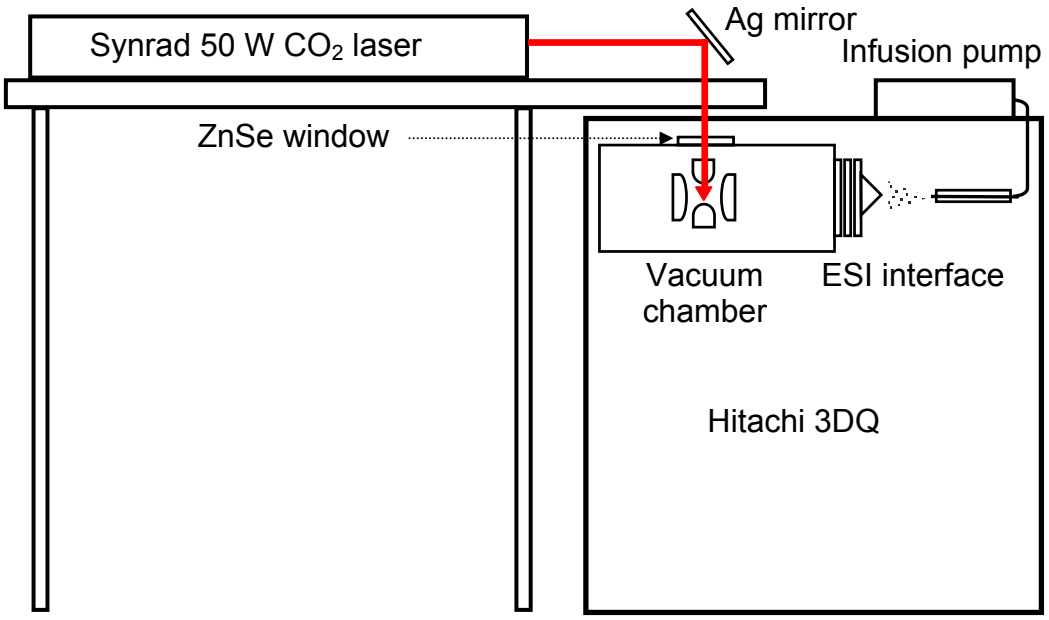


Figure 2.4: Schematic diagram of the Hitachi M-8000 QIT mass spectrometer modified for IRMPD.

In IRMPD experiments, ions isolated by DC/RF were activated by CO_2 laser irradiation. To allow detection of characteristic small fragment ions, the RF trapping voltage was adjusted to establish a low mass cutoff at m/z 70 during the ion activation period. After the dissociation period, fragment ions were detected by mass-selective axial instability ejection with axial modulation.

2.3 REFERENCES

(1) Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. *Anal. Chim. Acta* **1991**, *246*, 211-225.

(2) Sannes-Lowery, K.; Griffey, R. H.; Kruppa, G. H.; Speir, J. P.; Hofstadler, S. A. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1957-1961.

(3) Dunbar, R. C. *Int. J. Mass Spectrom.* **2000**, *200*, 571-589.

(4) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, *88*, 4451-4459.

(5) Cole, R. B., Ed. *Electrospray Ionization Mass Spectrometry*; 1st ed.; Wiley-Interscience: New York, 1997.

(6) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.

(7) March, R. E. *J. Mass Spectrom.* **1997**, *32*, 351-369.

(8) Hakansson, K.; Axelsson, J.; Palmblad, M.; Hakansson, P. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 210-217.

(9) Sannes-Lowery, K. A.; Hofstadler, S. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1-9.

Chapter 3: Comparison of SORI-CAD and MSAD for Top-Down Protein Analysis

3.1 INTRODUCTION

Mass spectrometry (MS) has played a critical role in the rapid advance of proteomics research.[1] Protein samples are often analyzed using a "bottom-up" approach, which typically involves one- or two-dimensional chromatography or electrophoresis, proteolytic digestion of the resulting fractions, and mass spectrometric fingerprinting of component peptides.[2-4] Tandem mass spectrometry is sometimes performed to provide a more thorough analysis of particularly complex protein mixtures.[5-7] However, the enzymatic digestion process central to the bottom-up approach not only complicates sample handling but also generates large quantities of data that can be difficult and time-consuming to process. In addition, valuable information about post-translational modifications to the sequence encoded by the genome can be lost because intact protein masses are not determined.

An alternative "top-down" strategy eliminates the digestion process and instead involves direct MS/MS analysis of intact proteins.[8-10] In this approach, primary sequence information derives exclusively from fragmentation patterns, so robust ion activation techniques are critical. The most effective methods of protein fragmentation include sustained off-resonance irradiation collision-activated dissociation (SORI-CAD),[11,12] nozzle-skimmer dissociation (NS),[13,14] infrared multiphoton dissociation (IRMPD),[15] and electron capture dissociation (ECD).[16-18] In a recent report CAD, IRMPD, and ECD were used in combination to provide backbone coverage sufficient to establish the primary

sequence(s) of several proteins using MS data alone.[19] In most applications, however, fragmentation data would be used to query a database of known sequences to identify any previously characterized proteins that may be present in a sample of unknown composition. In several studies the sequence tag construct identified by Mann et al.[20] has been shown to be an effective tool in this process.[21-24] In other recent reports it has been demonstrated that existing databases can be queried with unprocessed product ion data in conjunction with experimental molecular masses.[25,26] This approach provides flexibility in situations where sequence tags are either not observed or provide insufficient data for unambiguous identification, and online access to this search strategy has recently become available.[27]

To date top-down protein analysis has largely been restricted to infusion experiments, in which analyte concentrations are constant. In a variety of applications, such as LCMS or high-throughput protein array analysis techniques under current development at ORNL,[28] analytical signals are transient in nature. Fourier-transform ion cyclotron resonance (FTICR) platforms offer excellent resolution and mass range capabilities in such applications. However, certain dissociation methods available in FTICR instruments, such as SORI and ECD, possess relatively low duty cycles, and significant portions of a transient signal profile may be poorly sampled. IRMPD offers a relatively high duty cycle, and was successfully incorporated by Li et al. in the LC-FTMS analysis of a mixture of intact proteins.[29]

Another efficient but less well-characterized method of ion dissociation is multipole storage assisted dissociation (MSAD).[30] In this technique, space charge-mediated dissociation occurs in the linear RF-only multipoles commonly used to interface electrospray ionization with FTICR cells. Under certain conditions (extended accumulation times and/or larger dc offset voltages), ion activation and dissociation can

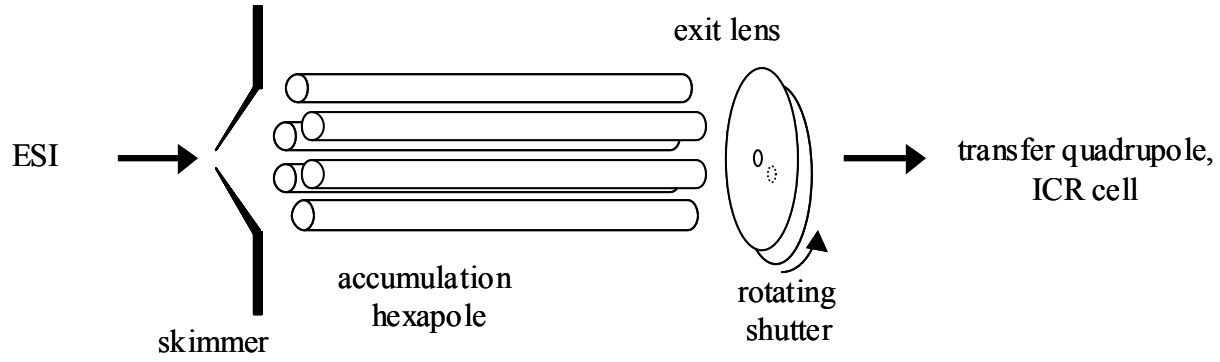
be induced during the ion accumulation interval in an inherently high-pressure region of the instrument, so no pump-down periods are required and the overall duty cycle is quite high. Linear ion traps are often used to interface electrospray ionization with FTICR cells, so MSAD can be readily performed on many FTICR instruments with no additional components. Sannes-Lowery et al.[31] and Hakansson et al.[32] demonstrated that MSAD is fundamentally a collisionally activated process, which implies that MSAD could in theory access the same fragmentation pathways as more established collisional activation methods like SORI. Hakansson et al.[32] compared MSAD and SORI for the 3 kD peptide melittin and found that the resulting spectra were indeed quite similar. It is not clear, however, how these two methods compare for larger proteins, which tend to produce broader charge state distributions upon electrospray ionization. In contrast with SORI, where individual charge states would typically be isolated and dissociated separately, MSAD activates the entire ensemble simultaneously. Contributions from distinct fragmentation pathways are therefore more likely in MSAD experiments, and because protein dissociation patterns are known to be charge state dependent[33-35]somewhat different fragmentation information could therefore result. For this reason a systematic comparison of SORI and MSAD was undertaken for a variety of 8-18 kDa proteins, with particular emphasis on the relative ability of these methods to provide fragmentation information that can be used in database search strategies for the positive identification of the parent proteins.

3.2 EXPERIMENTAL

Bovine ubiquitin and bovine β-lactoglobulin A were acquired from Sigma and used with no additional purification. Recombinant TNF-α (mouse and human), IL-2 (human), IL-3 (human), and IFN-γ (mouse, rat, and human) were acquired from

PeproTech Inc. (Rocky Hill, NJ) and reconstituted according to manufacturer's recommendations. Samples for mass spectrometry were prepared at formal concentrations of 5-10 μM in 50:50:1 (v/v/v) acetonitrile:water:acetic acid. Mouse and human IFN-γ were desalted using C_4 ZipTips (Millipore, Bedford, MA) prior to dilution.

Mass analysis was performed with a HiResESI Fourier-transform ion cyclotron resonance mass spectrometer (IonSpec, Lake Forest, CA) equipped with a 9.4T magnet (Cryomagnetics Inc., Oak Ridge, TN). Samples were introduced to an electrospray interface (Analytica of Branford, CT) by infusion at 1-3 μl/min. Ions were accumulated in an external hexapole situated between the skimmer cone on one end and an exit lens and mechanical shutter on the other (diagram, Figure 3.1). The pressure in this region of the instrument was typically around 2×10^{-5} Torr. At the end of the accumulation period, the exit lens voltage was dropped to zero and the mechanical shutter was pulsed open to allow ion transfer to the ICR cell. In SORI experiments, ion accumulation (typically 300 to 1200ms) was followed by a short period (130ms) for ion isolation, which was accomplished with a SWIFT pulse. Activation (1-4 V_{p-p}, 1s) was performed at a frequency 1 kHz lower than the parent ion cyclotron frequency in the presence of nitrogen, which was admitted with a pulsed valve to a transient pressure of 5×10^{-6} Torr. An 8-10 s pump-down delay was inserted prior to detection to produce overall scan functions of 12-15 s per transient. For MSAD experiments simultaneous ion accumulation/activation was usually performed for 500 to 4000 ms at a hexapole offset voltage of –8V (Fig. 3.1). This creates a deeper potential well than the standard offset setting (-3.5V) and therefore promotes ion fragmentation during the accumulation period.[36] Discrete isolation, activation, or pump-down delay times were not necessary, so overall scan functions for MSAD generally took 2 to 6 s per transient.



Scan Function Step		Time (ms)	Skimmer (V)	Hexapole DC offset (V)	Exit lens (V)	Shutter (V)
Ion accumulation:	Standard parameters	300-1200	+10	-3.5	+4.5	-20
	MSAD parameters	500-4000	+10	-8	+4.5	-20
Ion injection		160	+10	0	0	-20

Figure 3.1: Schematic diagram of the hexapole accumulation region of the HiResESI FTICR mass spectrometer.

Each spectrum was comprised of ten coadded transients acquired at 512K data points/transient, and external calibration was performed with ubiquitin; these conditions typically result in mass accuracy of ± 5 ppm and resolution of 150,000 (FWHM). Product ion spectra were deconvoluted to zero charge state with the IonSpec software deconvolution tool, and sequence tags were identified by manual inspection of the deconvoluted spectra. Product ion assignments were then made by comparison with theoretical isotopic distributions generated by the IonSpec exact mass calculator.

Database searches using the sequence tag strategy were performed using the TagIdent tool of the ExPASy molecular biology server.[37] For each protein the Swiss-PROT and TrEMBL databases were queried with experimentally determined molecular mass and sequence tag information. To facilitate identification of post-translationally modified proteins, a search tolerance of ± 500 Da was allowed for molecular mass data. The TagIdent tool restricts sequence tag entries to a maximum of six residues; in cases where longer tags were obtained, six residue subsequences containing a minimum number of ambiguous residues were chosen. All sequence tags were submitted in both the forward and reverse directions, and no keyword, species, or pI constraints were imposed.

3.3 RESULTS AND DISCUSSION

3.3.1 Fragmentation Patterns Determined by MSAD and SORI

Dissociation results obtained in MSAD experiments at various accumulation times were compared with data from SORI activation of 4-7 different charge states of each protein. In a few cases, the two dissociation methods gave very similar results. For other proteins, however, distinct differences between the MSAD and SORI data were

evident. Several factors that appear to create or contribute to spectral differences between MSAD and SORI spectra were identified.

Charge State Distribution: MSAD and SORI data for bovine ubiquitin are presented in Figure 3.2. Standard electrospray conditions (hexapole offset = -3.5 V) produced a charge state distribution dominated by the +10 to +12 charge states as shown in Fig. 3.2A. Under MSAD conditions (hexapole offset = -8V) an accumulation time of 1250 ms produced significant dissociation of the intact protein and gave a spectrum dominated by the y58, y24, and y18 ions (Fig. 3.2B, C). These results are qualitatively similar to data reported by Sannes-Lowery et al.,[30] who inferred that the dominance of just a few dissociation channels was observed because MSAD intrinsically selects for the lowest energy processes. However, direct comparison of MSAD and SORI data for this protein suggests that charge state dependent fragmentation may be another possible explanation. The major species in Fig. 3.2A (the +10 to +12 ions) gave relatively narrow product ion distributions in SORI experiments as illustrated by data for the +10 ion (Fig. 3.2D). This is consistent with previously reported results obtained by CAD in a quadrupole ion trap.[33] In contrast, the intermediate charge states (the +7 to +9 ions) gave a greater variety of fragments, as illustrated by the SORI spectrum of the +8 ion (Fig. 3.2E). These ions carry less overall charge, so proton mobility along the amide backbone is increased and a wider array of products results.[33] Fig. 3.2B indicates that the MSAD experiment consumed essentially the entire population of the +10 and +11 charge states, while detectable levels of the intermediate charge states (+7 to +9) survived. In addition, a comparison of Figs. 3.2C-E illustrates that MSAD produced results far more similar to SORI of the +10 ion than SORI of the +8 ion. These observations suggest that spectra generated by MSAD may simply reflect the different fragmentation tendencies of the dominant charge states. This is also consistent with the fact that MSAD of the charge

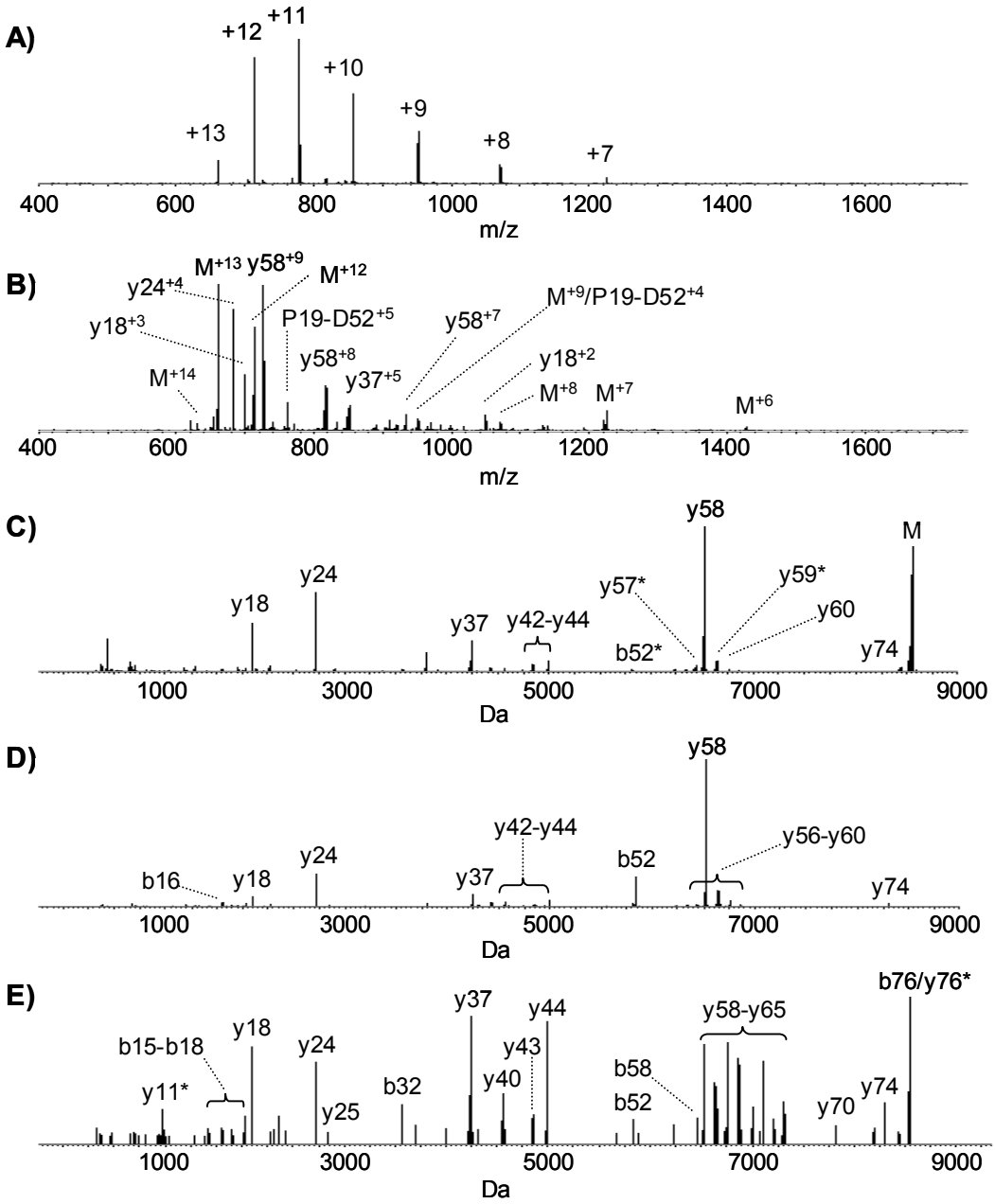


Figure 3.2: Dissociation data for bovine ubiquitin. A) Positive ion ESI mass spectrum, 750 ms accumulation, hexapole offset –3.5V. B) MSAD, 1250 ms accumulation, hexapole offset –8V. C) MSAD, 1250 ms accumulation, deconvoluted to zero charge state. The large peak for intact ubiquitin resulted from non-uniform actiationof the original charge state envelope (vide infra). D) SORI, +10 charge state (deconv). E) SORI, +8 charge state (deconv). Asterisks denote standard b/y fragments that have undergone loss of ammonia or water.

states in Fig. 3.2A gave less overall sequence coverage and less sequence tag information (see below) than collective SORI results for the +7 to +12 charge states. This apparent charge state dependence implies that in infusion experiments it may be possible to tune MSAD for particular fragmentation channels by manipulating parameters that affect the observed charge state distribution, such as the composition and pH of the electrospray solvent.

Radial Stratification: MSAD spectra acquired for recombinant mouse interferon-γ after 1000 and 2600 ms accumulation periods are shown in Figure 3.3A-C, and the SORI spectrum obtained from the +15 charge state is shown in Figure 3.3D. Similar N-terminal sequence tags (y114-y120 and y115-y121) dominate the product ion distributions in both Fig. 3.3C and D, which suggests that energy distribution in MSAD and SORI can be quite comparable. This is reflected in the fact that both activation methods produced approximately the same sequence coverage and sequence tag data for this particular protein.

In SORI experiments, the activation voltages required for 70% conversion of parent to backbone cleavage products followed the trend +17 < +16 < +15 < +14. However, Figure 3.3B shows that MSAD preferentially consumed the *lowest* charge states first. This behavior was also observed in MSAD of multiply charged peptides by Smith et al., who demonstrated with theoretical calculations that high m/z species tend to adopt larger radial orbits in linear ion traps.[36,38] This radial stratification results in greater average kinetic energies at the outer portions of the ion cloud. As accumulation time increases, increasing space charge within the linear trap causes the ion cloud to expand radially, increasing the likelihood that high m/z species (i.e. the lower charge states) will dissociate upon collision with background neutrals. Non-uniform activation can therefore supersede intrinsic trends in dissociation thresholds across a given charge state

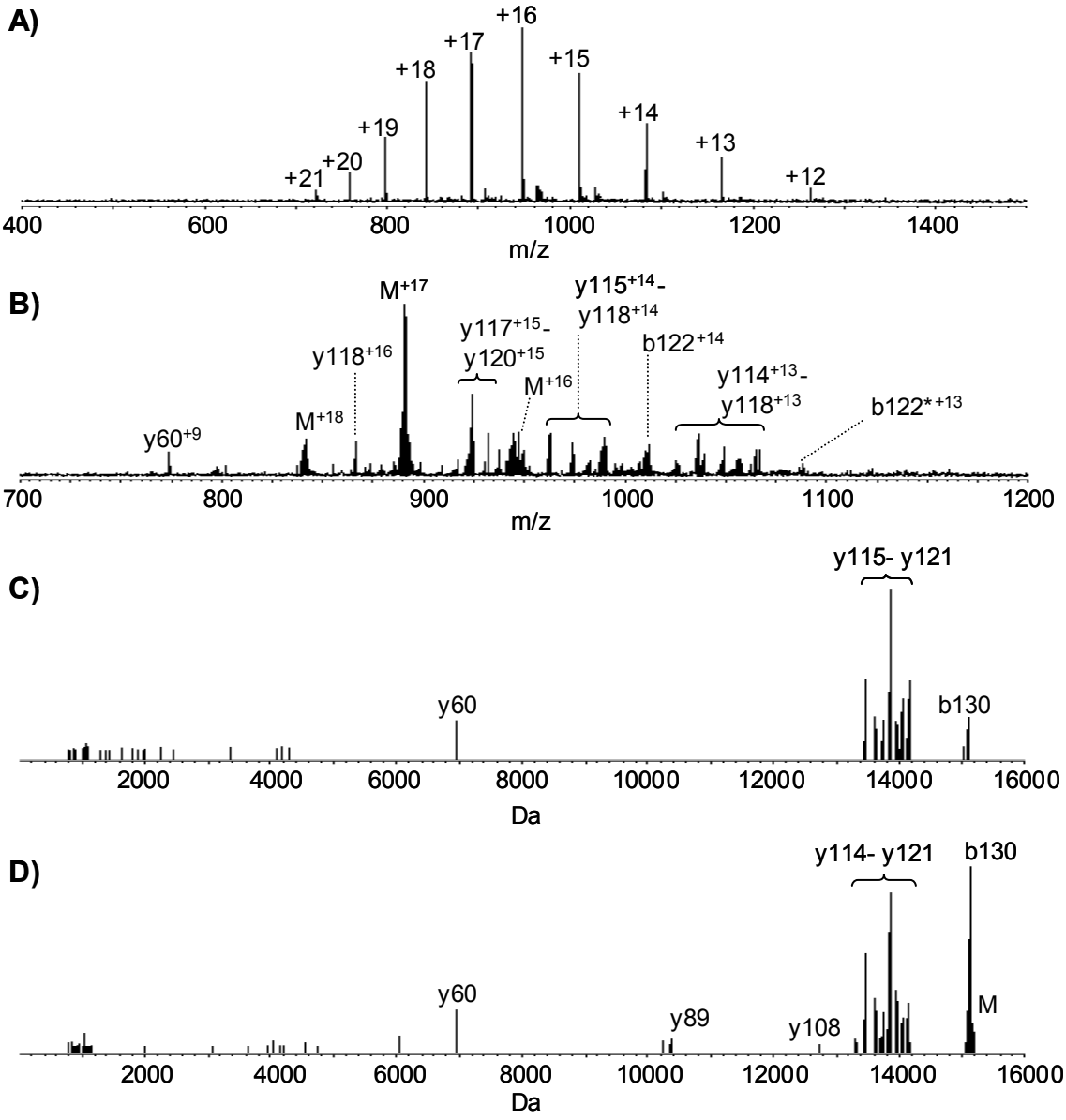


Figure 3.3: Dissociation data for mouse IFN-γ. A) Positive ion ESI mass spectrum, 1000 ms accumulation, hexapole offset –8V. Peaks corresponding to a minor sequence variant contaminating the sample are marked (*). B) MSAD, 2600 ms accumulation (deconv). D) SORI, +15 charge state (deconv). Asterisks denote standard b/y fragments that have undergone loss of ammonia or water.

distribution. MSAD spectra for most proteins investigated in this study showed some evidence of this phenomenon.

In the case of mouse IFN-γ, radial stratification had little impact on the fragmentation patterns observed by MSAD relative to those from SORI experiments. A distinct impact was observed in the dissociation data for other proteins, however. MSAD data acquired for horse apomyoglobin after accumulation periods of 750, 1000, and 2000 ms is presented in Figure 3.4. A comparison of Figs. 3.4A-C illustrates the tendency of MSAD to reduce overall charge density and compress ion signal into the low m/z region of the spectrum. At the 750 ms time point a broad charge state envelope was observed, but after 1000 ms the charge state distribution was much narrower and only a modest amount of dissociation is evident. This suggests that because of radial stratification within the hexapole, increasing space charge either propelled the lowest charge states to radii that exceed that of the conductance limit (exit lens, Fig. 3.1), or led to ejection rather than dissociation of the lower charge states. Among the species lost were the "intermediate" charge states (+12 to +14), which have been shown to provide the widest variety of fragment ions,[34] and as for ubiquitin the overall sequence coverage and sequence tag data retrieved for this protein by MSAD suffered as a result.

While radial stratification seems to have been responsible for premature loss of a several parent ions for apomyoglobin, product ion information seems to have been compromised for human TNF-α. SORI and MSAD data for this protein are presented in Figures 3.5 and 3.6, respectively. SORI of the +14 charge state (Fig. 3.5A and B) produced a spectrum dominated by the y138 and y136 species, in addition to a 7-residue sequence tag comprised of the b149-156 ions. MSAD consumed most of the +14 charge state after 3750 ms (Figs. 3.6B and C), but only a portion of the sequence tag was

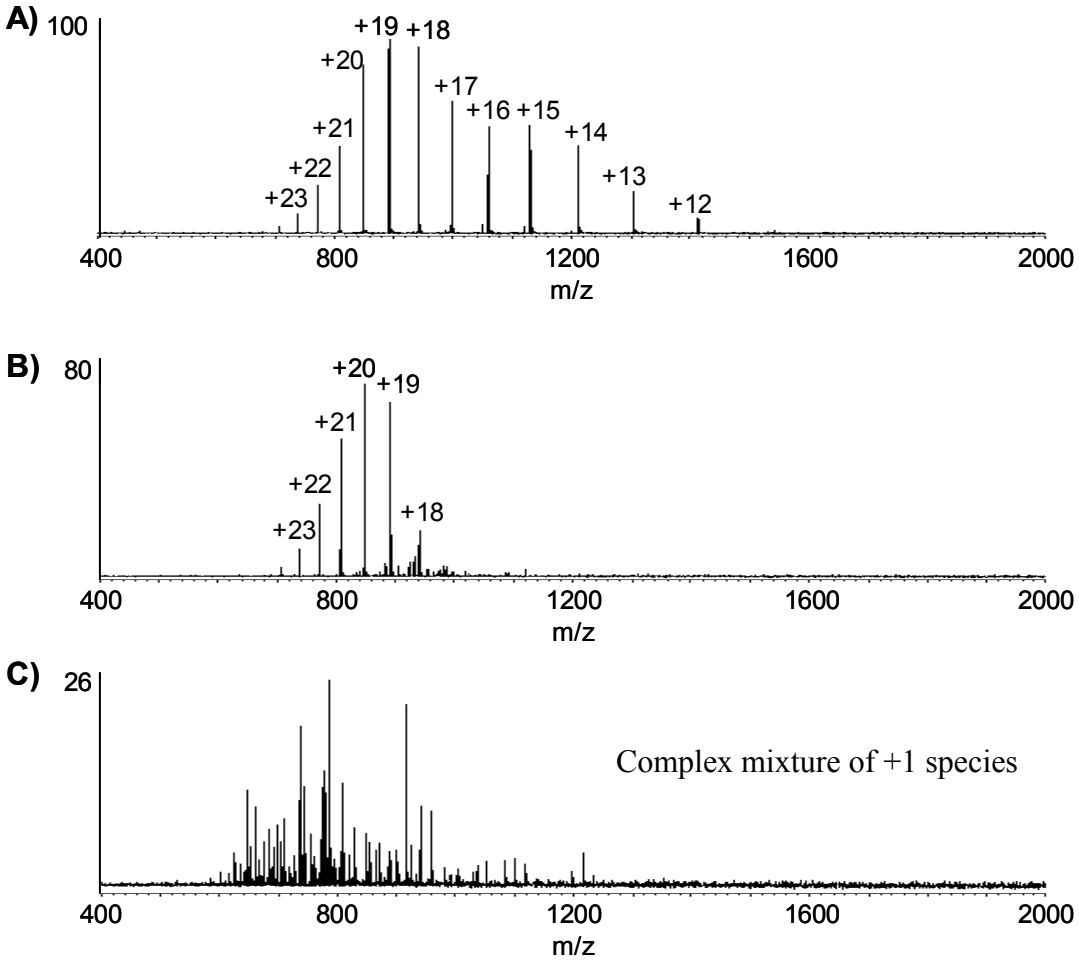


Figure 3.4: Spectra acquired for horse apomyoglobin under MSAD conditions (hexapole offset = -8 V). A) 750 ms accumulation, B) 1000 ms accumulation, C) 2000 ms accumulation.

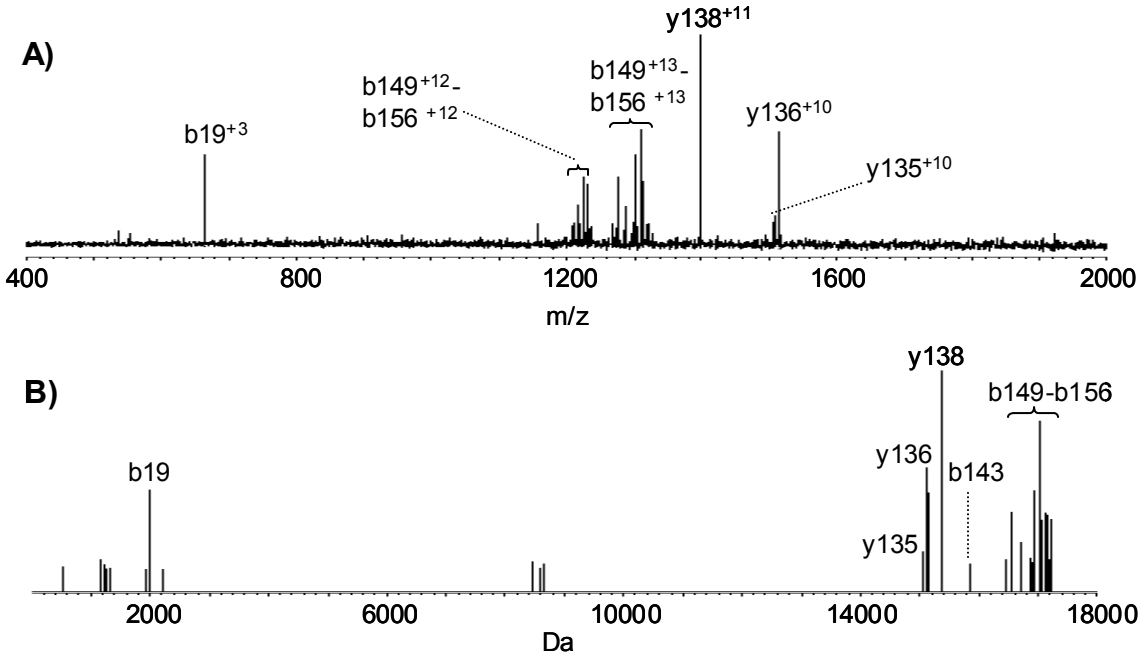


Figure 3.5: Dissociation data for human TNF-α. A) SORI, +14 charge state, B) SORI, +14 charge state (deconv.), C) Positive ion ESI spectrum, 1200 ms, hexapole offset –8V, D) 3750 ms accum., E) 3750 ms accum. (deconv.) Asterisks denote standard b/y fragments that have undergone loss of water or ammonia.

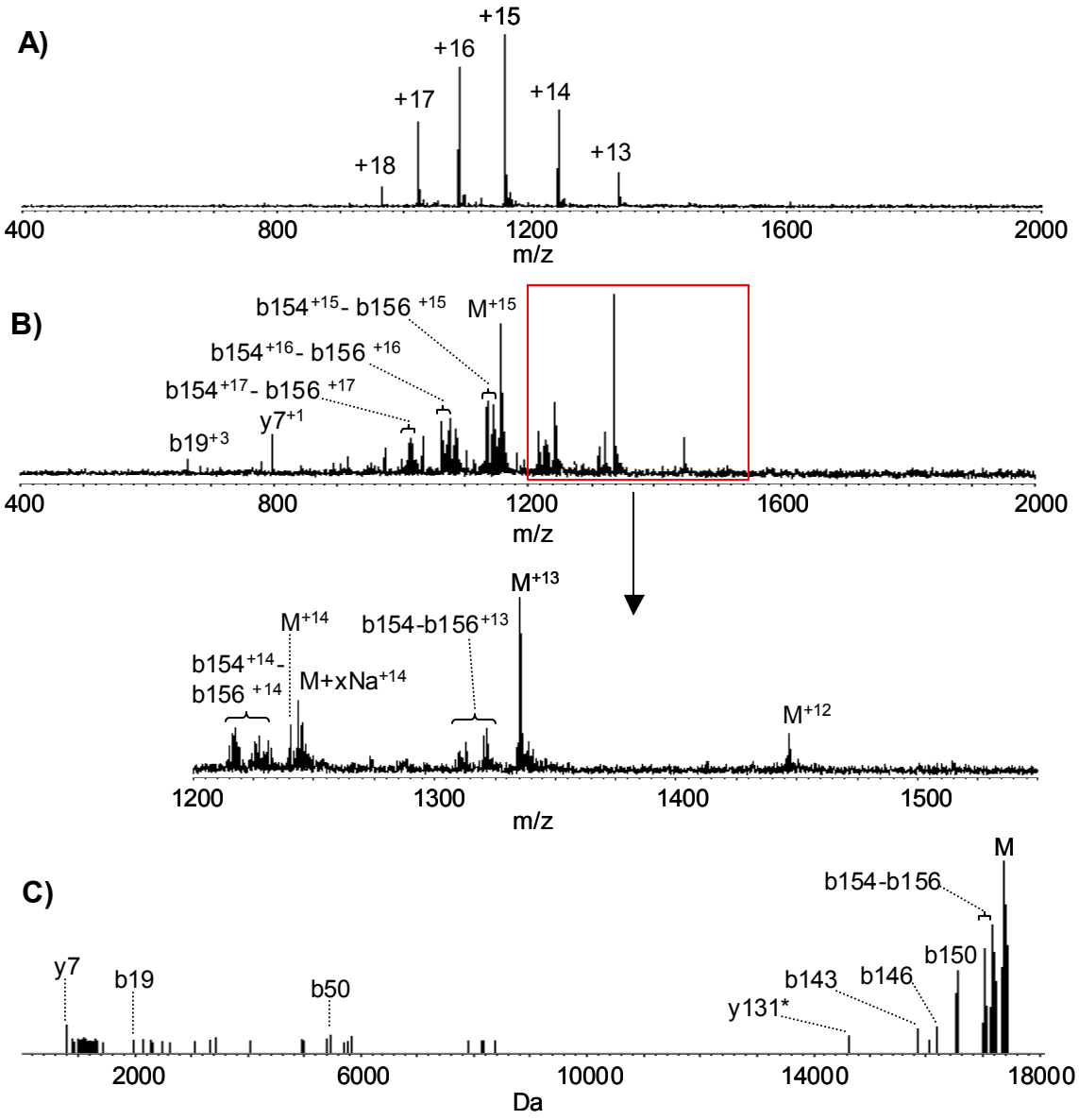


Figure 3.6: Dissociation data for human TNF-α. A) Positive ion ESI spectrum, 1200 ms, hexapole offset –8V, B) 3750 ms accum., C) 3750 ms accum. (deconv.) Asterisks denote standard b/y fragments that have undergone loss of water or ammonia.

recovered and the y138 and y136 ions were missing entirely. These fragments constitute a major dissociation pathway so it seems unlikely that differences in energy distribution in the SORI and MSAD experiments would account for this result. However, in Fig. 3.5A it can be seen that the missing y ions form in the high m/z region of the spectrum. It is possible that in the MSAD experiment these fragment ions adopt radii that exceed the radius of the conductance limit in the optics region and are therefore not transferred to the cell for mass analysis. It is also possible that at high space charge the effective potential well is too shallow to trap high m/z fragments. Either of these explanations may also account for the fact that the b149 and b151-155 ions that help define the sequence tag in Fig. 3.5B do not appear in Figs. 3.6B and C.

Non-Resonant Activation: SORI and MSAD data for mouse TNF-α are presented in Figure 3.7 and 3.8. SORI of the +12 charge state produced a wide variety of fragment ions, including an extended sequence tag comprised of the b142-b152 and b154-b156 ions (Figs. 3.7A-C) along with several complementary ions (y6-y8 and y11-y12) at low intensity. At moderate accumulation times (Figs. 3.8B-C), MSAD produced only some of the b ions comprising the sequence tag, and those ions that do appear derive from dissociation of the +13 charge state. The +12 charge state was completely consumed, however, and comparison with Fig. 3.7A suggests that, again, trapping and/or transmission of high m/z ions may have been impaired under conditions of high space charge and the b144-151^{+11} and b142-146^{+10} ion series were therefore not observed. Another difference between the Figs. 3.7B and 3.8C is that the complementary y ions were observed in higher signal intensity as an intact sequence tag in the MSAD experiment. The improved response for these species by MSAD could be due to relatively poor trapping of low m/z species in the SORI experiment, relatively efficient trapping of low m/z species in the accumulation hexapole, or both.

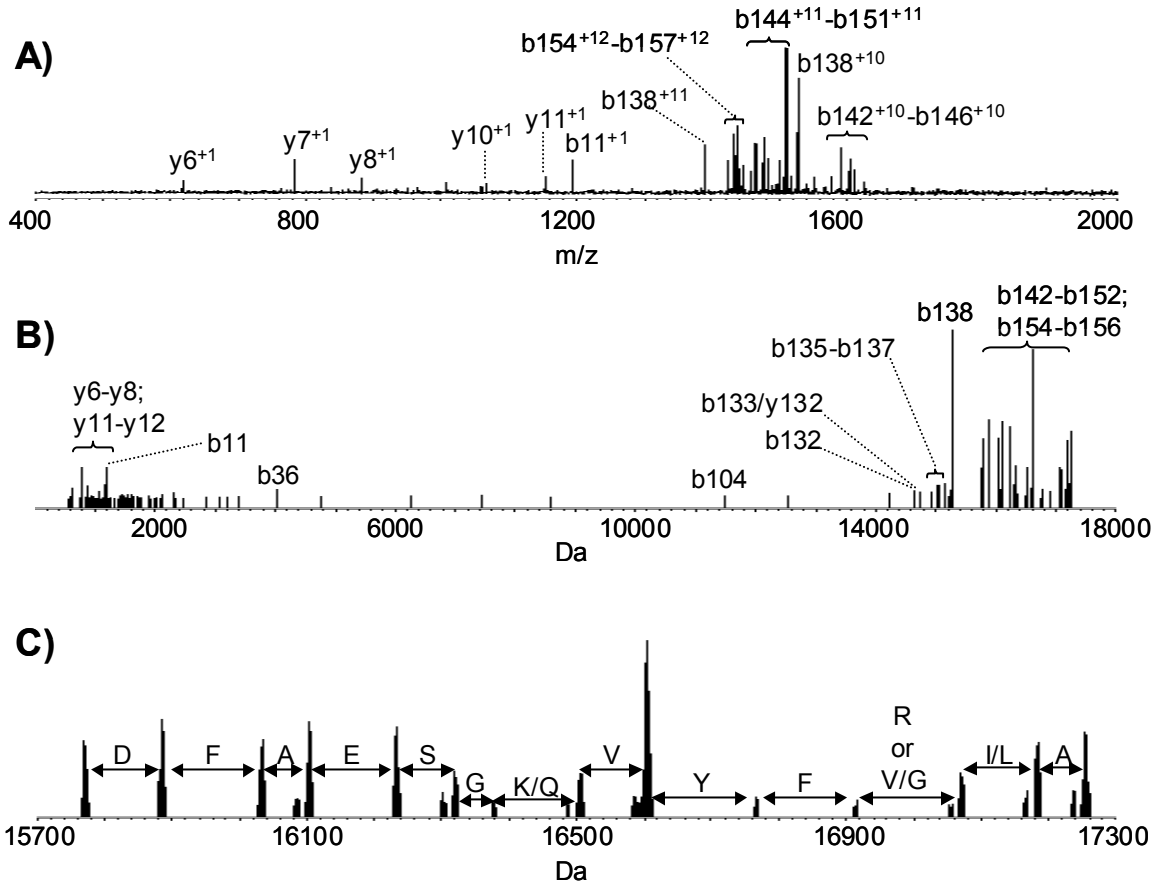


Figure 3.7: SORI data for mouse TNF-α. A) SORI, +12 charge state, B) +12 charge state (deconv.), C) Expansion of sequence tag (deconv.). Asterisks denote standard b/y fragments that have undergone loss of ammonia or water.

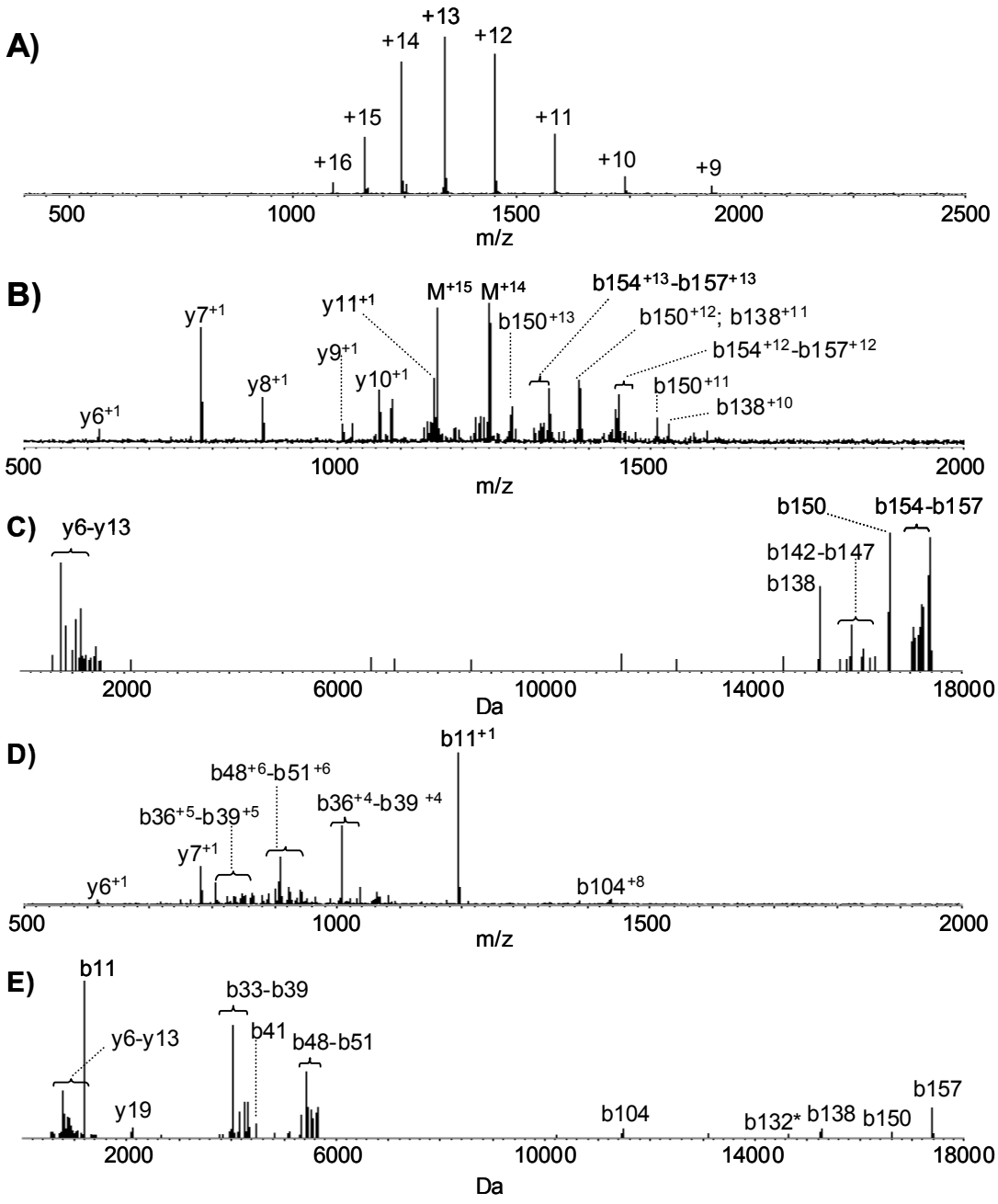


Figure 3.8: MSAD data for mouse TNF-α. A) Positive ion ESI mass spectrum, 1000 ms accumulation, hexapole offset = -8V, B) MSAD, 2000 ms accumulation, C) MSAD, 2000 ms accumulation (deconv.), D) MSAD, 3000 ms accumulation, and E) MSAD, 3000 ms accumulation (deconv.). Asterisks denote standard b/y fragments that have undergone loss of ammonia or water.

At longer MSAD times (Figs. 3.8D-E), a number of smaller ions defining two sequence tags (b48-b51 and b33-b39, corresponding to the sequences PVV(I/L) and A(N/D)A(I/L)(I/L)A) were observed. These fragments form by secondary dissociation of the larger b ions found in Figs. 3.8B and C, and correspond to the loss of a portion of the amide backbone constrained by this protein's single disulfide bond (C70-C101). Only two of the ions comprising these additional sequence tags were observed in SORI experiments. Because of the non-resonant nature of activation, MSAD provided better access to such secondary dissociation processes for this protein, and as a result MSAD produced greater sequence coverage and more sequence tags than SORI in this case.

3.3.2 Comparison of Database Search Information Generated by SORI and MSAD

Tables 3.1 and 3.2 summarize the dissociation data obtained by SORI and MSAD, respectively. For each activation method, the observed sequence tags and corresponding fragment ion assignments are listed, along with the number of unique proteins ("hits") retrieved from the ExPASy molecular biology server for each sequence tag. The tables illustrate that all proteins in this study produced sequence tags by SORI, and that at least some sequence tag data can also be retrieved by MSAD. For mouse IFN-γ, human IL-3, and human IL-2, the sequence tag data provided by the two activation methods was nearly the same. For ubiquitin and apomyoglobin, however, it appears that charge state availability imposed by either the electrospray conditions or the effects of radial stratification restricted the level of fragmentation data observed by MSAD and therefore prevented identification of the parent protein. For human TNF-α, mouse TNF-α, and β-lactoglobulin A, the effects of radial stratification in MSAD disrupted the primary sequence tag seen in SORI experiments but also allowed trapping of a complementary

Table 3.1: Fragmentation Data, SORI Experiments

Protein (MW)	Sequence Tags				Preferred Site Fragments Observed
	Residues	Fragment Ions[1]	Charge State(s)	Search Results	
Bovine ubiquitin (8.6 kD)	T(I/L)T(I/L)EVE VEPSD KE KE	y58-y65 y55-y60 y42-y44 y12-y14	+8 +10 +9, +10 +9	13 hits[2] --- --- ---	7 of 8
Human IL-3 (15.1 kD)	(I/L)S(I/L)A(I/L)	b128*-b133*	+13	75 hits	1 of 15
Mouse IFN-γ (15.1 kD)	ES(I/L)NNY	y115-y121	+15, +16	1 hit[3] (correct)	0 of 8
Human IL-2 (15.5 kD)	TFA(K/Q)S(I/L)(I/L)	b123-b130	+14	0 hits[4]	1 of 8
Horse apomyoglobin	(N/I/L)SDG(E/M)	y147-153	+17	59 hits	9 of 12
Human IFN-γ (16.9 kD)	(K/Q)VMAE(I/L)S (K/Q)D	y22-y29 y141-y143	+19 +19	6 hits[5] ---	4 of 11
Human TNF-α (17.4 kD)	VYFG(I/L)(I/L)A	b149-b156	+14	16 hits[2]	5 of 15
Mouse TNF-α (17.4 kD)	DFAESG(K/Q)VYF(R/GV/VG)(I/L)A	b142-152, b154-156	+12	14 hits[2]	4 of 17
β-Lactoglobulin A (18.3 kD)	AMAASDIS(I/L)(I/L)DA(K/Q)SA	y125-y140	+16	11 hits[2]	4 of 19

[1]Asterisks denote standard b/y fragments that have undergone loss of ammonia or water.
[2]All hits were homologs of the correct protein.
[3]Search constraints: sequence tag, MW ± 750 Da.
[4]The correct database entry was not retrieved because of an amino acid substitution (C → A) within the sequence tag.
[5]Most hits were homologs of the correct protein.

Table 3.2: Fragmentation Data, MSAD Experiments

Protein (MW)	Sequence Tags				Preferred Site Fragments Observed
	Residues	Fragment Ions[1]	Accum. Time (ms)	Search Results	
Bovine ubiquitin (8.6 kD)	EP (K/Q)E (K/Q)ES	y57*-59* y42-44 y11-14	1500 1500 2500	> 3000 --- ---	7 of 8
Human IL-3 (15.1 kD)	SNA(I/L)F	b128*-133*	2500	0 hits[6]	3 of 15
Mouse IFN-γ (15.1 kD)	(S/F/Y)(I/L)NNYF	y114-120	3000	1 hit[3] (correct)	2 of 8
Human IL-2 (15.5 kD)	TFA(K/Q)S(I/L)(I/L)	b123-130	2750	0 hits[4]	1 of 8
Horse apomyoglobin	Y(K/Q)	y6-8	1500	---	9 of 12
Human IFN-γ (16.9 kD)	(K/Q)V S(I/L)	internal fragments y26-28	2600 2600	--- ---	5 of 11
Human TNF-α (17.4 kD)	(I/L)AN (I/L)VV	b154-157 b47-50	3500 4500	>3500 >3000	6 of 15
Mouse TNF-α (17.4 kD)	AESG(K/Q)VY PVV(I/L) A(N/D)A(I/L)(I/L)A	y6-13 b48-52 b33-39	2000 2500 3000	19 hits[2] --- 15 hits[5]	7 of 17
β-Lactoglobulin A (18.3 kD)	AMAAS G(I/L)D	b22-27 b8-11	2000 3000	16 hits[5] ---	2 of 19

[1]Asterisks denote standard b/y fragments that have undergone loss of ammonia or water.
[2]All hits were homologs of the correct protein.
[3]Search constraints: sequence tag, MW ± 750 Da.
[4]The correct database entry was not retrieved because of an amino acid substitution (C → A) within the sequence tag.
[5]Most hits were homologs of the correct protein.
[6]Insufficient sampling produced distorted isotopic distributions for some fragments comprising the sequence tag. One residue (N) was therefore incorrectly identified and the correct database entry could not be retrieved

ion series that provided information about the same portion of the amide backbone. Additional sequence tag information was also generated through secondary dissociation processes. For mouse TNF-α and β-lactoglobulin A positive identification of the parent protein was therefore still possible.

Tables 3.1 and 3.2 also list the number of "preferred" sites[34] (N-terminal to prolines and C-terminal to aspartic acids) where backbone cleavage was observed for each protein. This data illustrates that although sequence tag data was in some cases compromised by MSAD, cleavage at "preferred" sites was observed at least as often in MSAD as in SORI experiments. Search algorithms that utilize raw mass data provide alternatives to the sequence tag approach,[25,26] and typically weight fragments derived from cleavage at "preferred" sites more heavily than other fragments. MSAD data might therefore be utilized more effectively within the framework of this newer type of search strategy. In addition, because lower charge states can be lost during MSAD, algorithms that weight preferred cleavages associated with these parent ions less heavily than those associated with higher charge states could be advantageous.

3.4 CONCLUSIONS

MSAD spectra of intact proteins are influenced by several factors. Data for bovine ubiquitin suggests that because dissociation patterns are charge state dependent, the charge state distribution made available by the ionization conditions may dictate the range of fragment ions that can be generated during the experiment. In addition, conditions of high space charge within the hexapole impair transmission and/or trapping of high m/z species, which can result in loss of important precursor and product ions. And finally, the non-resonant nature of activation in MSAD can provide access to secondary dissociation processes that are not available by SORI. Because of these

considerations, the two activation methods do not always provide the same fragmentation information, and MSAD is somewhat less reliable for generating sequence tag data. However, it appears that in general MSAD samples "preferred" cleavage processes (i.e. those occurring at D and P residues) just as well as SORI, which implies that MSAD data may be somewhat more compatible with search algorithms that utilize unprocessed fragment ion masses.

Because MSAD scan functions do not require extended pump-down delays or discrete ion isolation and activation periods, individual MSAD experiments in this study required one-sixth to one-half the time needed for SORI experiments. MSAD further reduced overall analysis time in that, unlike SORI, it was not necessary to select and dissociate an array of different charge states in order to characterize a given protein. It remains to be seen, however, if these advantages can be exploited in the analysis of transient signals. The non-resonant nature of MSAD could provide rapid, high-throughput analysis of multiple analytes, but it may also complicate data interpretation for complex mixtures. In addition, the charge density dependence of MSAD[30,31]potentially limits its effective dynamic range. Low concentration samples would require long accumulation times and/or large hexapole offset potentials, which might adversely affect sensitivity. Another relevant consideration is that because concentration varies across a transient peak (from low to high to low), the extent of fragmentation observed across that peak would probably vary as well. Judicious choice of MSAD parameters, including accumulation time and hexapole offset potential, would be necessary to optimize the overall figures of merit, and multiplexed scan functions might prove beneficial in this regard. Additional experiments will therefore be required to fully evaluate MSAD for applications such as surface sampling or LCMS.

3.5 REFERENCES

(1) Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001**, *101*, 269-295.

(2) Patterson, S. D. *Anal. Biochem.* **1994**, *221*, 1-15.

(3) Jensen, P. K.; Pasa-Tolic, L.; Peden, K. K.; Martinovic, S.; Lipton, M. S.; Anderson, G. A.; Tolic, N.; Wong, K. K.; Smith, R. D. *Electrophoresis* **2000**, *21*, 1372-1380.

(4) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R., 3rd *Nat. Biotechnol.* **1999**, *17*, 676-682.

(5) Hunt, D. F.; Yates, J. R., III; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U S A* **1986**, *83*, 6233-6237.

(6) Kassel, D. B.; Musselman, B. D.; Smith, J. A. *Anal. Chem.* **1991**, *63*, 1091-1097.

(7) Wilm, M.; Shevchenko, A.; Houthaeve, T.; Breit, S.; Schweigerer, L.; Fotsis, T.; Mann, M. *Nature* **1996**, *379*, 466-469.

(8) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 806-812.

(9) Reid, G. E.; McLuckey, S. A. *J. Mass Spectrom.* **2002**, *37*, 663-675.

(10) VerBerkmoes, N. C.; Bundy, J. L.; Hauser, L.; Asano, K. G.; Razumovskaya, J.; Larimer, F.; Hettich, R. L.; Stephenson, J. L., Jr. *J. Proteome Res.* **2002**, *1*, 239-252.

(11) Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. *Anal. Chim. Acta* **1991**, *246*, 211-225.

(12) Senko, M. W.; Speir, J. P.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2801-2808.

(13) Senko, M. W.; Beu, S. C.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 415-418.

(14) Smith, R. D.; Loo, J. A.; Barinaga, C. J.; Edmonds, C. G.; Udseth, H. R. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 53-65.

(15) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

(16) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

(17) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563-573.

(18) McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 245-249.

(19) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U S A* **2000**, *97*, 10313-10317.

(20) Mann, M.; Wilm, M. *Anal. Chem.* **1994**, *66*, 4390-4399.

(21) Mortz, E.; O'Connor, P. B.; Roepstorff, P.; Kelleher, N. L.; Wood, T. D.; McLafferty, F. W.; Mann, M. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8264-8267.

(22) Cargile, B. J.; McLuckey, S. A.; Stephenson, J. L. *Anal. Chem.* **2001**, *73*, 1277-1285.

(23) Demirev, P. A.; Ramirez, J.; Fenselau, C. *Anal. Chem.* **2001**, *73*, 5725-5731.

(24) Nemeth-Cawley, J. F.; Rouse, J. C. *J. Mass Spectrom.* **2002**, *37*, 270-282.

(25) Meng, F.; Cargile, B. J.; Miller, L. M.; Forbes, A. J.; Johnson, J. R.; Kelleher, N. L. *Nat. Biotechnol.* **2001**, *19*, 952-957.

(26) Reid, G. E.; Shang, H.; Hogan, J. M.; Lee, G. U.; McLuckey, S. A. *J. Am. Chem. Soc.* **2002**, *124*, 7353-7362.

(27) In https://prosightptm.scs.uiuc.edu; The University of Illinois at Urbana-Champaign.

(28) Van Berkel, G. J.; Kennel, S. J.; Doktycz, M. J.; Ford, M. J.; Sanchez, A. D.; Quirke, J. M. E. In *Proceedings of the 51st ASMS Conference on Mass Spectrometry and Allied Topics*: Montreal, Canada, 2003.

(29) Li, W.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Anal. Chem.* **1999**, *71*, 4397-4402.

(30) Sannes-Lowery, K.; Griffey, R. H.; Kruppa, G. H.; Speir, J. P.; Hofstadler, S. A. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1957-1961.

(31) Sannes-Lowery, K. A.; Hofstadler, S. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1-9.

(32) Hakansson, K.; Axelsson, J.; Palmblad, M.; Hakansson, P. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 210-217.

(33) Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal. Chem.* **2001**, *73*, 3274-3281.

(34) Newton, K. A.; Chrisman, P. A.; Reid, G. E.; Wells, J. M.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2001**, *212*, 359-376.

(35) Hogan, J. M.; McLuckey, S. A. *J. Mass Spectrom.* **2003**, *38*, 245-256.

(36) Belov, M. E.; Gorshkov, M. V.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1312-1319.

(37) In http://us.expasy.org/tools/tagident.html; Swiss Institute of Bioinformatics/North Carolina Supercomputing Center.

(38) Tolmachev, A. V.; Udseth, H. R.; Smith, R. D. *Int. J. Mass Spectrom.* **2003**, *222*, 155-174.

Chapter 4: Collision-Activated Dissociation and Infrared Multiphoton Dissociation of Oligonucleotides in a Quadrupole Ion Trap

4.1 INTRODUCTION

Infrared multiphoton dissociation (IRMPD) offers an promising alternative to collision-activated dissociation (CAD) in quadrupole ion trap mass spectrometry.[1-11] In the quadrupole ion trap (QIT), efficient activation by CAD requires an elevated RF trapping voltage, which means that fragments with m/z values less that about one-third of the parent ion mass may not be trapped. In contrast, energy deposition in IRMPD can be manipulated independently of the RF level, and as a result low mass fragments not seen in CAD spectra can often be observed in IRMPD experiments. IRMPD is usually performed with CO_2 lasers, which provide 10.6 μm radiation that is absorbed by a variety of organic ions regardless of their m/z value. Activation is therefore far less specific in IRMPD than in CAD, where resonant waveforms are supplied to deposit energy only into a narrow population of parent ions but not the ensuing fragment ions. As a result IRMPD can generally provide greater access to secondary dissociation processes and therefore supply structural information not readily retrieved by CAD without extensive MS^n sequences. Because of these characteristics IRMPD has proved an effective and advantageous tool in the analysis of an array of biologically active small molecules including macrolide,[1] β-lactam,[2] tetracycline, [3] and aminoglycoside antibiotics.[4]

IRMPD has become an established technique for the analysis of peptides,[12-16] intact proteins,[17-21] and oligonucleotides[17,22-28] in FTICR mass spectrometers. In the quadrupole ion trap, however, the utility of IRMPD as a tool in the study of biopolymers

is less well characterized. A few reports describing IRMPD of peptides and proteins in the QIT have appeared recently[9,10,29,30] but IRMPD of oligonucleotides is limited to a few examples.[7,11] A systematic comparison of CAD and IRMPD for the analysis of five- to forty-residue oligonucleotides has therefore been performed, and the relative utility of the two techniques for oligonucleotide sequencing has been evaluated.

4.2 EXPERIMENTAL

HPLC-purified oligonucleotides were obtained from TriLink Biotechnologies Inc. (San Diego, CA) in the ammonium salt form. Samples for analysis by negative ion electrospray contained 10 μM oligonucleotide, 5 mM piperidine, and 5 mM imidazole in a 70:30 mixture (v/v) of acetonitrile and water. Samples for analysis by positive ion electrospray contained 20 μM oligonucleotide in a 50:50:1 mixture (v/v) of methanol, water, and acetic acid. Test solutions were infused (0.5 – 1 μl/min) through a custom electrospray source consisting of a fused silica sample capillary, stainless steel sheath needle, and nitrogen assist gas.

All experiments were performed with a modified Hitachi M-8000 quadrupole ion trap mass spectrometer[31] operating in negative ion mode. To allow IRMPD, a hole was cut in the vacuum chamber directly over the ion trap, and a custom flange bearing a 1.5" ZnSe window (II-VI Inc.) was bolted over it. In addition, a 6-mm hole was drilled radially through one side of the ring electrode. The beam from a continuous 50W CO_2 laser (Model 48-5, Synrad Inc., Mukilteo WA) was directed through both the ZnSe window and the hole in the ring electrode with a copper mirror (II-VI Inc.). To trigger the laser, TTL pulses generated by the M-8000 software during the ion activation period were routed to a laser control module (Synrad Model UC-2000). For ion isolation a temporary DC offset was applied to the trap in lieu of the standard filtered noise field

(FNF) waveforms supplied by the M-8000. DC inputs provided by two external power supplies were applied to the endcap electrodes in a dipolar fashion through the RF board by means of an in-house switching circuit controlled by software-generated TTL signals. This DC/RF approach provided greater ejection efficiency for high m/z ions than the RF/FNF strategy employed on the standard M-8000.[32] The instrument pressure during mass analysis was nominally 5 x 10^{-5} Torr for all experiments.

4.3 RESULTS AND DISCUSSION

In this study we have examined the CAD and IRMPD mass spectra for a variety of oligonucleotides in order to compare the diagnostic information obtained from each. The presence or absence of specific types of fragment ions is noted, as well as the degree of sequence coverage afforded by each activation method.

CAD and IRMPD spectra for the –3 charge state of the 5-mer d(CGTTC) and the –4 charge state of the 8-mer d(CGAGCTCG) are shown in Figure 4.1 as representative examples illustrating the quality of the data and the types of fragments that can be observed. In both CAD spectra (Fig. 4.1A and C), the dominant ions correspond to loss of nucleobases from the intact deoxyribose/phosphate backbone (M-B ions in the McLuckey nomenclature).[33] These ions relate no sequence information but do decompose to informative backbone cleavage products.[33] For the 5-mer (Fig. 4.1A), thymine is lost in preference to cytosine and guanine. This behavior is unusual, however, and is observed because this particular sequence does not contain adenine. Results more typical of mixed base oligonucleotides are observed for the 8-mer (Fig. 4.1C), where loss of adenine predominates.[34-39] The backbone cleavage products in Figs. 4.1A and C all

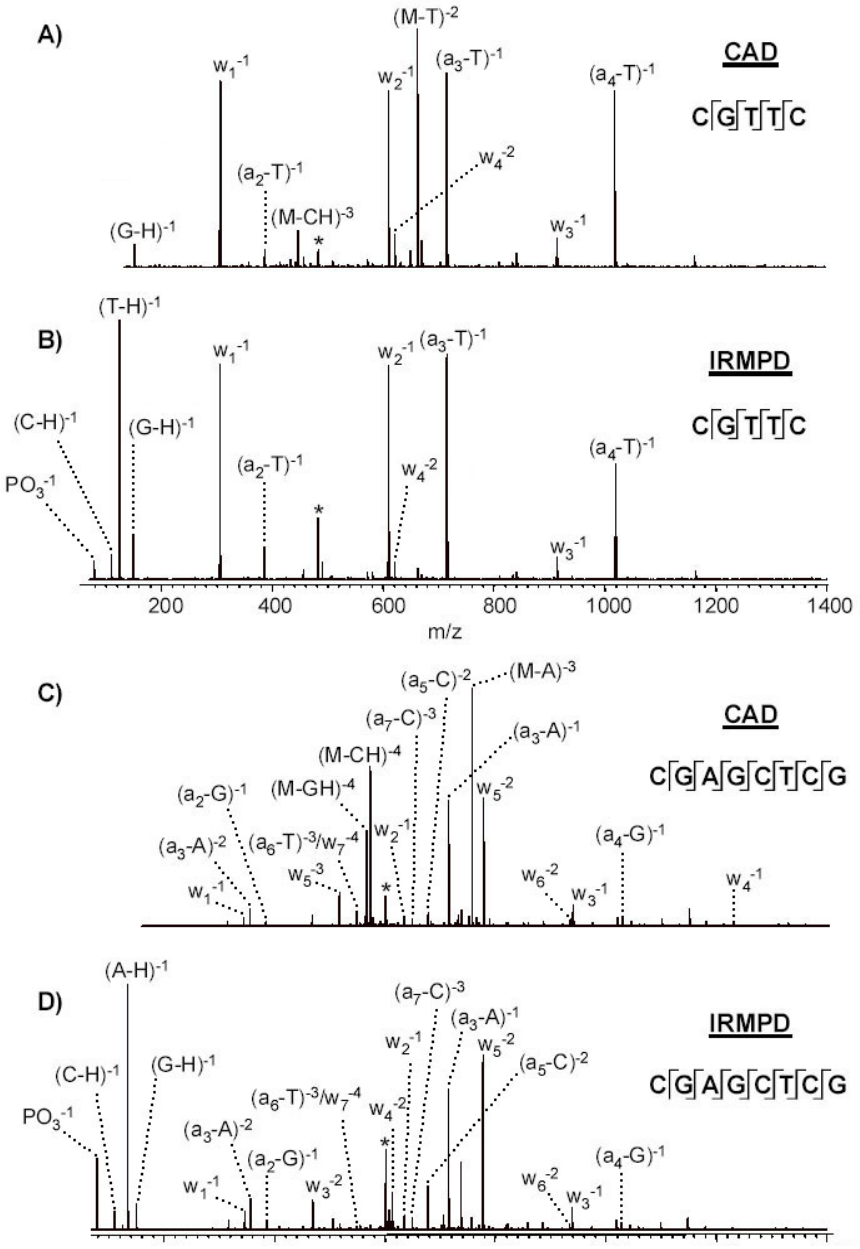


Figure 4.1: MS/MS data for d(CGTTC) and d(CGAGCTCG). A) CAD (165 mV/50 ms) and B) IRMPD (34 W/50 ms) of the 5-mer, -3 charge state. C) CAD (215 mV/50 ms) and D) IRMPD (27 W/50 ms) of the 8-mer, -4 charge state. Parent ions are marked with asterisks.

belong to the w or a-B series as expected from previous reports.[17,33,34] For both of these small oligonucleotides, CAD provided complete sequence coverage and a full array of complementary fragment ions.

In the corresponding IRMPD spectra (Fig. 4.1B and D) the major backbone cleavage products are also w and a-B ions. In contrast with the CAD spectra, however, the M-B ions are greatly reduced in intensity. This is no doubt a result of non-resonant activation in IRMPD and implies that dissociation thresholds for conversion of M-B ions to backbone cleavage products are in general lower than those for generation of M-B species from of the parent ion. In addition, because the low mass cutoff can be manipulated independently of ion activation in IRMPD, several low mass ions that could not be trapped in CAD experiments are visible in Fig. 4.1B and D. These include the PO_3^- ion, diagnostic of the phosphate backbone, as well as the deprotonated base anions (such as $(T\text{-}H)^{-1}$ in Fig. 4.1B) whose formation is only implied in the CAD spectra by the charges on complementary ion pairs (ex. w_2^{-1} and $(a_3 - T)^{-1}$ in Fig. 4.1A). Both IRMPD experiments gave complete sequence coverage in addition to a full array of complementary sequence ions on the same time scale used in the CAD experiments.

In the quadrupole ion trap, the efficiency of photoactivation is somewhat compromised by the relatively high operating pressures required for effective trapping and resonant ejection. In several previous studies of IRMPD in the QIT, efficient photoactivation has required either unusually low instrument pressure or an unusually high temperature. [1,8,9] The spectra in Figs. 4.1B and D, however, were acquired at normal instrument pressure and temperature, as were IRMPD spectra for all other oligos in this study. This was possible because of the strong absorbance of 10.6 μm radiation by the backbone phosphodiester moieties.[7,17] We have recently demonstrated that because of the presence of these chromophores, phosphorylated peptides dissociate much more

readily by IRMPD in a quadrupole ion trap than the corrresponding non-phosphorylated sequences.[29] For oligonucleotides the most important cleavage processes occur at the phosphate linkages themselves, so in this work clear correlations between absorbance cross-section and dissociation thresholds could not be made.

In IRMPD experiments, energy deposition and dissociation efficiency are moderated by adjusting either the laser power or the irradiation time. To illustrate the impact of these parameters on oligonucleotide dissociation, energy-variable IRMPD data for the –4 charge state of d(CGAGCTCG) are shown in Figure 4.2. Increasing either exposure time (Fig. 4.2A) or laser power (Fig. 4.2B) increases the proportion of backbone cleavage products relative to the parent ion. Most fragment ions reached a relatively constant relative signal intensity, which suggests that secondary dissociation of sequence ions is minimal although activation in IRMPD is non-resonant. This may provide indirect evidence of a relationship between absorbance and dissociation, since the fragments are smaller than the parent and therefore contain fewer (chromophoric) phosphate groups. It is also potentially significant because destruction of sequence ions and loss of sequence information is a disadvantage sometimes associated with IRMPD of biomolecules in ICR studies.[16]

Dissociation spectra for the –4 and –5 charge states of the 10-mer d(ATGCTACGAG) are shown in Figure 4.3. Results for CAD of the –4 parent ion (Fig. 4.3A) conform well to literature results.[40] In addition, comparison of Fig. 4.3A and C illustrates that base loss patterns depend on the charge state of the parent ion.[41] For the –4 charge state the nucleobases were eliminated as neutrals, while for the –5 charge state ionic base loss predominated because of greater Coulombic repulsion in the higher charge state.[41] In the IRMPD spectra (Fig. 4.3B and D), signal intensities for M-B ions were 8-residue sequences (Fig. 4.1B and D). For this 10-mer both CAD and IRMPD gave

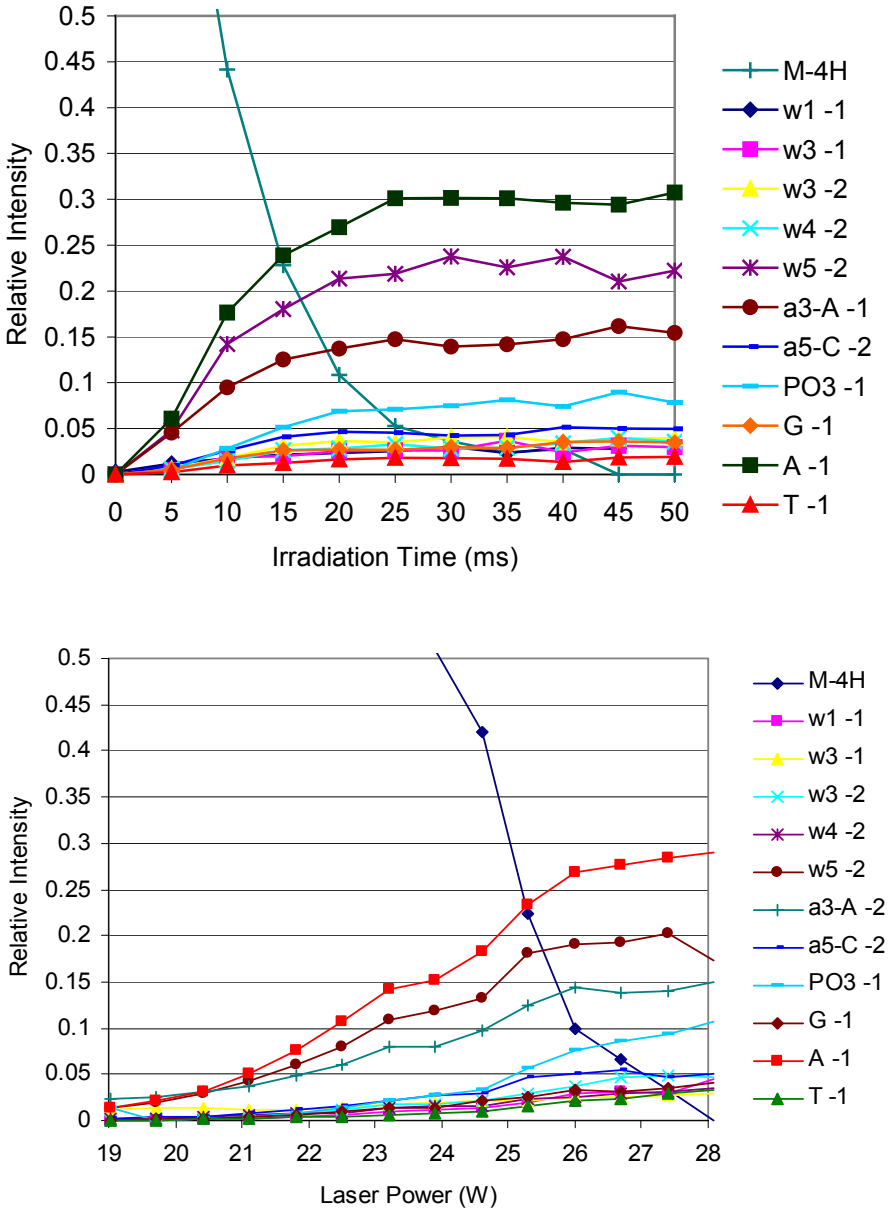


Figure 4.2: Energy-resolved IRMPD of d(CGAGCTCG), -4 charge state. A) Relative signal intensity vs. irradiation time, 15.5 W laser power. B) Relative signal intensity vs. laser power, 50ms irradiation.

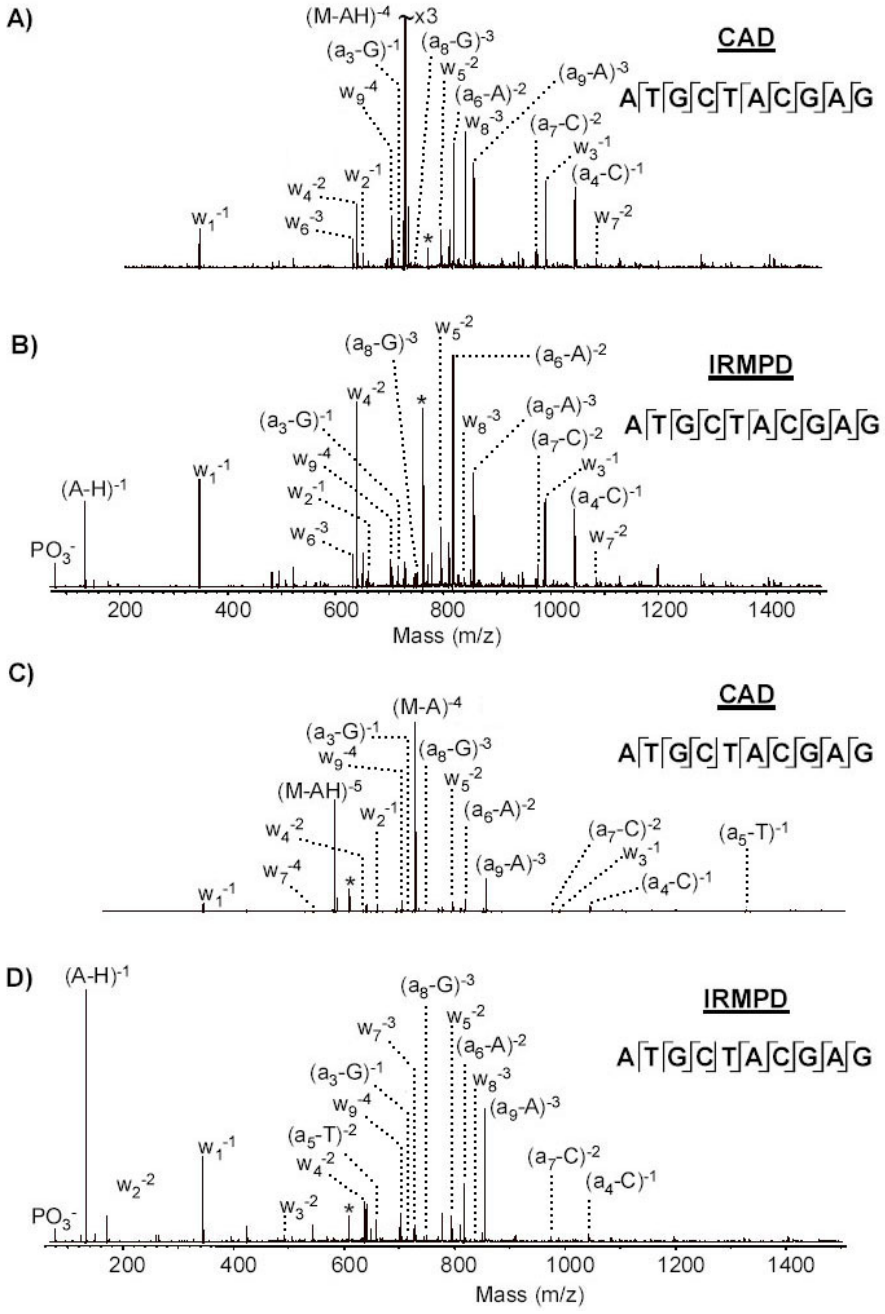


Figure 4.3: MS/MS data for d(ATGCTACGAG). CAD (240 mV/50 ms) and B) IRMPD (22.5 W/50 ms) of the –4 charge state. C) CAD (192 mV/50 ms) and D) IRMPD (21 W/ 50 ms) of the –5 charge state. Parent ions are marked with asterisks.

quite low but phosphate and deprotonated base ions were visible, as noted for the 5- and complete sequence coverage, and complementary fragments were observed for almost every sequence ion. It is also noteworthy that, in comparing CAD spectra in Figs. 4.1 and 4.3, some evidence of size discrimination in the activation process is apparent because the relative signal intensity for M-B ions is greater for the 10-mer than for the smaller oligonucleotides.

To further explore this possible trend for size discrimination in collisional activation, CAD and IRMPD spectra were acquired and compared for longer sequences. CAD and IRMPD spectra for the –10 charge state of the 20-mer d(GATCCTAGCTAGCTAGGATC) are shown in Figure 4.4. The size discrimination trend observed on going from 8 to 10 residues did seem to perpetuate in the CAD experiment (Fig. 4.4A). From Fig. 4.4B it is clear that extensive dissociation was obtained for this 6 kDa analyte by IRMPD. The charge state of most fragments cannot be determined from low-resolution QIT data, so approximately 10-30% of the sequence ion assignments are inherently somewhat ambiguous. However, confidence in many assignments was improved in both the CAD and IRMPD spectra because many fragments occurred in more than one charge state, and also because complementary w or a-B ions were also observed. Given this, the array of fragments obtained in the CAD and IRMPD experiments appears to be quite similar, and the sequence coverage is the same. The yield of sequence ions and the overall sequence coverage obtained by CAD and IRMPD for several other charge states of this oligonucleotide are given in Table 4.1, and these data show that in general the two activation methods provided similar information.

Dissociation data for the –14 charge state of the 40-mer oligonucleotide d(GACTACAAGTATGCGACGATGATTCTAGCTTACGTAGCCA) are given in

Charge State	Activation Method	w Ions[1]	a-B Ions[1]	Interresidue Positions Cleaved
-6	CAD	17	15	19
	IRMPD	19	18	19
-7	CAD	19	16	19
	IRMPD	19	17	19
-8	CAD	17	16	19
	IRMPD	18	16	19
-9	CAD	17	17	19
	IRMPD	18	16	19
-10	CAD	18	17	19
	IRMPD	16	18	19

[1]Sequence ions formed in multiple charge states counted only once.

Table 4.1: MS/MS data for d(GATCCTAGCTAGCTAGGATC).

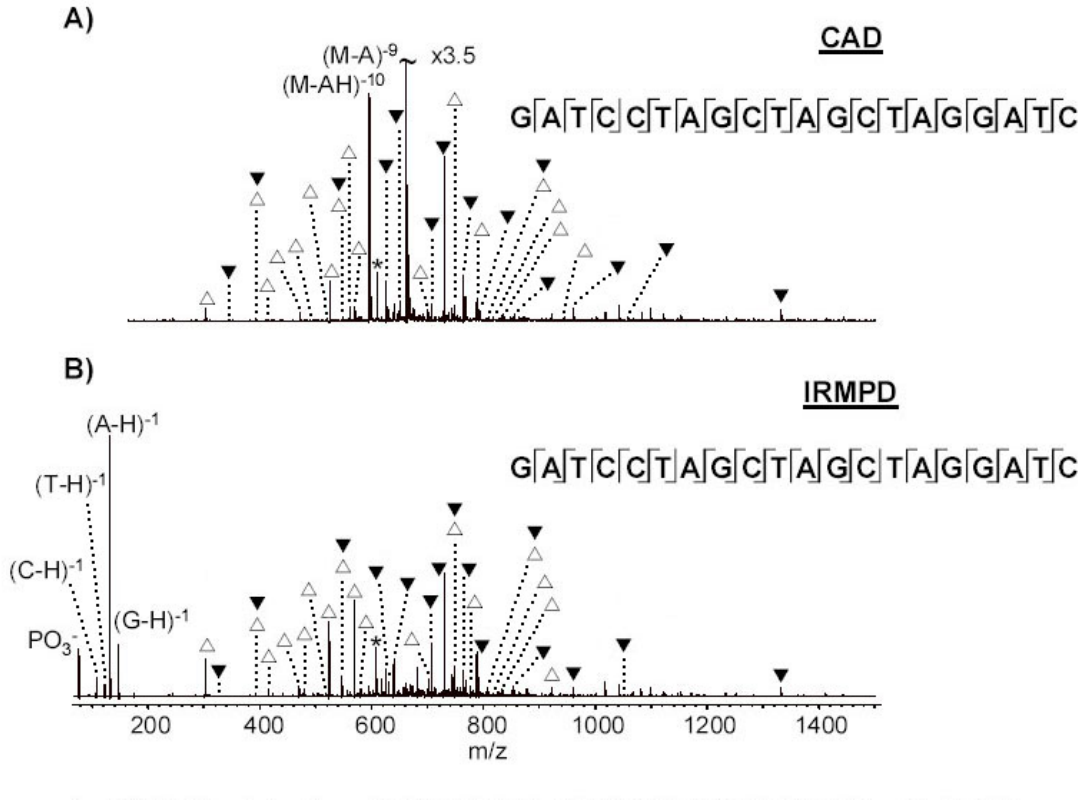


Figure 4.4: MS/MS data for d(GATCCTAGCTAGCTAGGATC). A) CAD (240 mV/50 ms) and B) IRMPD (17.6 W/50 ms) of the –10 charge state. Labeling scheme: * = parent ion, ▼ = a-B ion, △ = w ion.

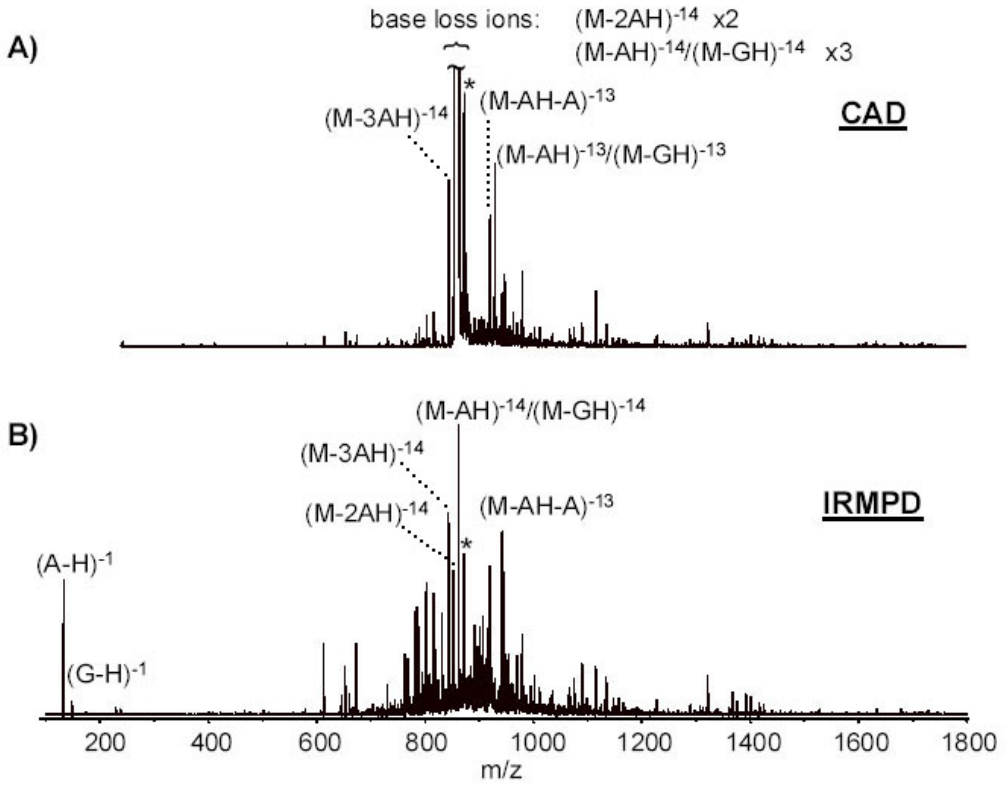


Figure 4.5: MS/MS data for d(GACTACAAAGTATGCGACGATGAGGCTAGCTT-ACGTAGCCA), -14 charge state. A) CAD 210 mV/50 ms, B) IRMPD, 14.8 W/50 ms. Parent ions are marked with asterisks.

Figure 4.5. (Because of the density of fragment ions, only peaks corresponding to the parent ion and M-B fragments are labeled.) For sequences of this length, only 60-70% of sequence ion assignments are unique, but the apparent sequence coverage is nearly the same for both CAD and IRMPD. In both spectra it is apparent that for oligonucleotides of this length loss of multiple base moieties can compete with backbone cleavage. Even the uninterpreted data demonstrate that IRMPD can dissociate large species (12kD) as effectively as CAD.

Because free nucleobase ions can be observed so readily by IRMPD, it seemed that IRMPD might facilitate the identification of modified residues. To explore this possibility, the 20-mer d(GATCC(N^6-MeA)GCTTACGGTACCA) was synthesized and analyzed by both CAD and IRMPD. The resulting spectra obtained from the –8 charge state are shown in Figure 4.6. The mass of the modified base at position 6, N^6-methyladenine, differs from that of guanine by only 2 amu (149.2 amu vs. 151.1 amu). The $(M\text{-}MeAH)^{-8}$ and $(M\text{-}GH)^{-8}$ ions for this 20-mer are therefore nearly isobaric, differing in mass by only 0.25 mass/charge units. CAD of the –8 charge state (Fig.4. 6A) gave a product at m/z 742.0 that, without prior knowledge of the sequence, could reasonably be assigned as an $(M\text{-}GH)^{-8}$ ion (theoretical m/z 741.5). Similarly, because most of the sequence ions in Fig. 4.6A are multiply charged, the majority could be assigned incorrectly on the basis of a sequence containing guanine rather than N^6-MeA at position 6. Evidence for the modification is subtle in the CAD experiment, and additional experiments would likely be required in order to confirm the presence and identity of the methylated base.

In contrast, the spectrum obtained by IRMPD clearly shows a peak at m/z 148 in addition to signals for the anions of the standard bases A, C, G, and T (Fig. 4.6B). This signal provides direct evidence of the modified base along with complete sequence

coverage in a single experiment. The ion at m/z 148 in the IRMPD spectrum was isolated and subjected to a second stage of activation, and the resulting MS^3 spectrum was used to confirm its identity based on a comparison with fragmentation patterns generated from authentic adenine and N^6-methyladenine (Figure 4.7). Deprotonated N^6-methyladenine dissociates by loss of H_2 or $^{\cdot}CH_3$, pathways identical to those observed for m/z 148 in Fig. 4.6B. The backbone phosphate groups of oligonucleotides are highly acidic and deprotonate readily, so oligonucleotides are most commonly analyzed in negative ion mode. Positively charged species can also be generated,[42-48] however, so a brief comparison of CAD and IRMPD for dissociation of positive charge states was also undertaken. Dissociation data for the +3 and +4 charge states of the 10-mer d(GACTACAAGT) are shown in Figure 4.8. As reported in previous CAD studies,[43-46] the major sequence ions are still w and a-B species, although they tend to occur in lower relative signal intensity in the positive mode. Consistent with previous reports, the 5'-terminal base dissociated preferentially to leave prominent GH^+ and w_9^{+n} ions in all spectra.[46] The CAD spectrum of the +3 charge state (Fig. 4.8A) does differ somewhat from a published spectrum acquired for this sequence in a triple quadrupole.[44] A number of low mass ions are missing, which is an effect of the low mass cutoff invariably associated with CAD in the quadrupole ion trap. In addition, no cleavage was observed 3' to position four, occupied by thymine. Selectivity against cleavage at T sites relative to A, G, and C sites has been noted in several previous studies conducted in quadrupole ion traps.[37-40] The fact that cleavage was observed at T4 in a triple quadrupole but not in the present study is at least in part a reflection of the slightly lower energy regime of the QIT. This notion is also supported by the fact that the M-B ions are far more prominent in Figure 8A than in the reference spectrum.

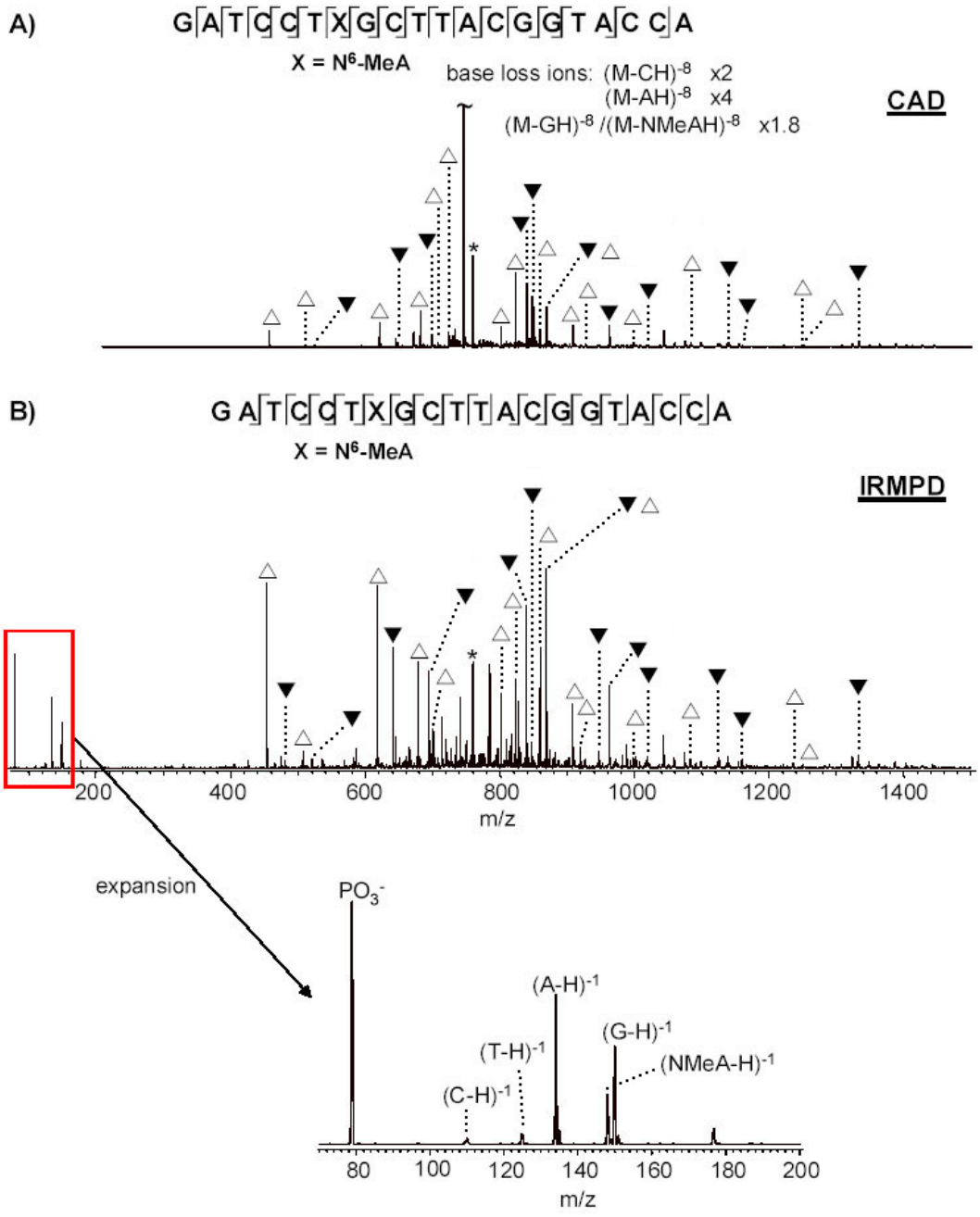


Figure 4.6: MS/MS data for d(GATCCG(N^6-MeA)GCTTACGGTACCA). A) CAD (215 mV/50 ms) and B) IRMPD (19.7 W/ 50 ms) of the –8 charge state. Labeling scheme: * = parent ion, ▼ = a-B ion, △ = w ion.

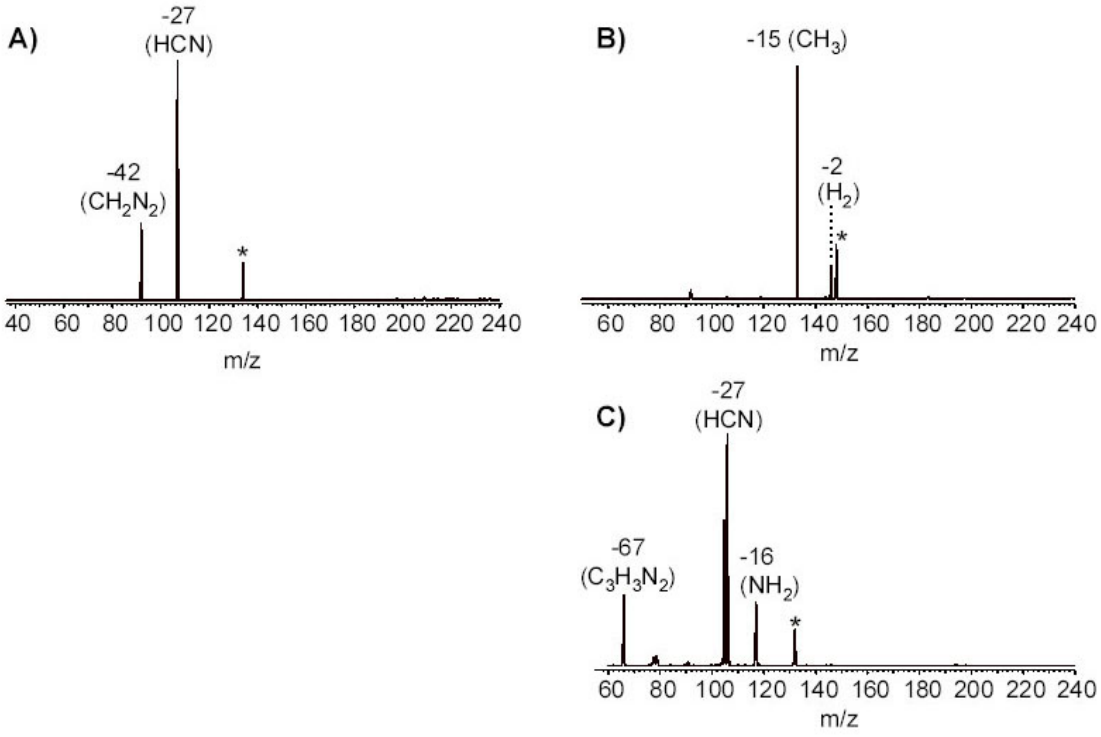


Figure 4.7: MS/MS data for deprotonated adenine and N^6-methyladenine. A) CAD of adenine $(M - H)^{-1}$, 94 mV/50 ms. B) CAD of N^6-methyladenine $(M-H)^{-1}$, 94 mV/50 ms. C) MS^3 of deprotonated N^6-methyladenine, m/z 148 → 133 → ?, 94 mV, 106 mV/50 ms. Parent ions are marked with asterisks.

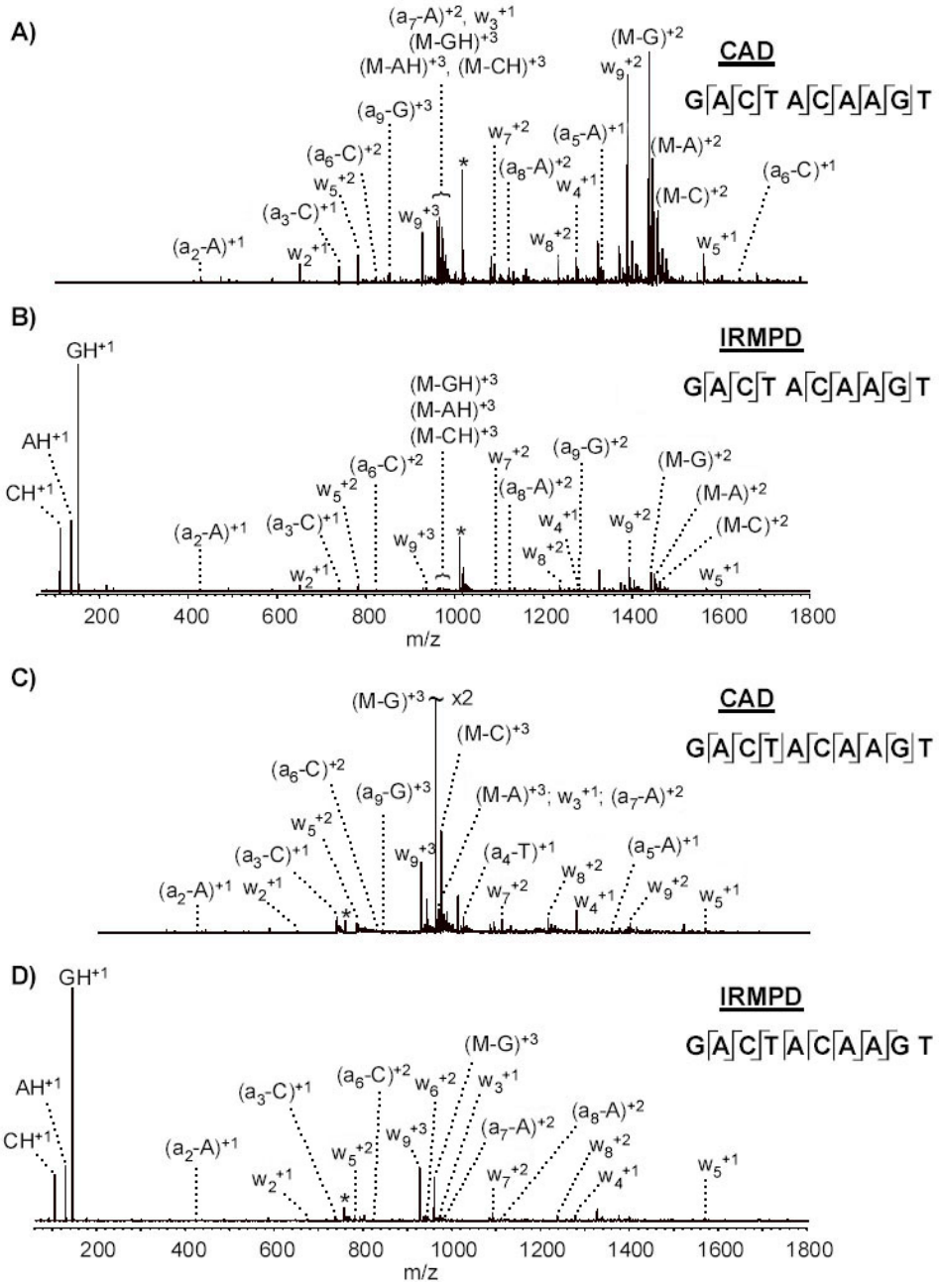


Figure 4.8: MS/MS data for d(GACTACAAGT). A) CAD (245 mV/50 ms) and B) IRMPD (18.3 W/50 ms) of the +3 charge state. C) CAD (180 mV/50 ms) and D) IRMPD (16.2 W/50 ms) of the +4 charge state.

Comparing the CAD and IRMPD spectra in Fig. 4.8 shows that the effects of the low mass cutoff and non-resonant activation noted for IRMPD experiments conducted in negative ion mode (trapping of base ions and other low mass species, consumption of M-B ions) were also evident in positive ion mode, and as in negative mode sequence coverage was essentially the same using either activation technique. The charge state dependence of base loss remained consistent as well: neutral base loss occurred to a greater extent for the lower charge state (Fig. 4.8A vs. C). However, the parent species that tend to lose bases in ionic form carried a lower overall charge in positive than in negative ion mode. The –4 charge state of this oligonucleotide lost neutral bases upon either CAD or IRMPD (data not shown), while the +4 charge state lost protonated bases (Fig. 4.8C and D). This likely occurred because of the high proton affinity of the purine and pyrimidine moieties.

4.4 CONCLUSIONS

IRMPD of deprotonated and protonated oligonucleotides (five to forty residues) has been performed in a quadrupole ion trap at normal operating pressure and temperature using moderate exposure times and laser powers. In general, IRMPD and CAD spectra of any given parent ion contain comparable information, indicating that IRMPD is a viable alternative to CAD for oligonucleotide analysis in this instrument platform. However, structurally uninformative M-B ions that dominate CAD spectra are usually found much lower signal intensity in IRMPD spectra because of the non-resonant nature of photoactivation. Also, phosphate and nucleobase ions can be observed directly in IRMPD experiments because the low mass cutoff can be set to trap small fragment ions. Because of this, analysis of some sequences containing modified bases can be facilitated by use of IRMPD.

4.4 REFERENCES

(1) Colorado, A.; Shen, J. X.; Brodbelt, J. *Anal. Chem.* **1996**, *68*, 4033-4043.

(2) Goolsby, B. J.; Brodbelt, J. S. *J. Mass Spectrom.* **1998**, *33*, 705-712.

(3) Vartanian, V. H.; Goolsby, B.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1089-1098.

(4) Goolsby, B. J.; Brodbelt, J. S. *J. Mass Spectrom.* **2000**, *35*, 1011-1024.

(5) Crowe, M. C.; Brodbelt, J. S.; Goolsby, B. J.; Hergenrother, P. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 630-649.

(6) Stephenson, J. L., Jr.; Booth, M. M.; Shalosky, J. A.; Eyler, J. R.; Yost, R. A. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 886-893.

(7) Stephenson, J. L., Jr.; Booth, M. M.; Boue, S. M.; Eyler, J. R.; Yost, R. A. *ACS Symp. Ser.* **1996**, *619*, 512-564.

(8) Boue, S. M.; Stephenson, J. L., Jr.; Yost, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1391-1397.

(9) Payne, A. H.; Glish, G. L. *Anal. Chem.* **2001**, *73*, 3542-3548.

(10) Hashimoto, Y.; Hasegawa, H.; Yoshinari, K.; Waki, I. *Anal. Chem.* **2003**, *75*, 420-425.

(11) Drader, J. J.; Hannis, J. C.; Hofstadler, S. A. *Anal. Chem.* **2003**, *75*, 3669-3674.

(12) Jockusch, R. A.; Paech, K.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 3188-3196.

(13) Masselon, C.; Anderson, G. A.; Harkewicz, R.; Bruce, J. E.; Pasa-Tolic, L.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 1918-1924.

(14) Flora, J. W.; Muddiman, D. C. *Anal. Chem.* **2001**, *73*, 3305-3311.

(15) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.

(16) Li, L.; Masselon, C. D.; Anderson, G. A.; Pasa-Tolic, L.; Lee, S. W.; Shen, Y.; Zhao, R.; Lipton, M. S.; Conrads, T. P.; Tolic, N.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 3312-3322.

(17) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

(18) Mortz, E.; O'Connor, P. B.; Roepstorff, P.; Kelleher, N. L.; Wood, T. D.; McLafferty, F. W.; Mann, M. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8264-8267.

(19) Li, W.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Anal. Chem.* **1999**, *71*, 4397-4402.

(20) Meng, F.; Cargile, B. J.; Miller, L. M.; Forbes, A. J.; Johnson, J. R.; Kelleher, N. L. *Nat. Biotechnol.* **2001**, *19*, 952-957.

(21) Meng, F.; Cargile, B. J.; Patrie, S. M.; Johnson, J. R.; McLoughlin, S. M.; Kelleher, N. L. *Anal. Chem.* **2002**, *74*, 2923-2929.

(22) Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893-4897.

(23) Little, D. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **1995**, *117*, 6783-6784.

(24) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **1996**, *118*, 9352-9359.

(25) Hofstadler, S. A.; Griffey, R. H.; Pasa-Tolic, L.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1400-1404.

(26) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 2067-2070.

(27) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 945-951.

(28) Sannes-Lowery, K. A.; Hofstadler, S. A. *J. Am. Soc.Mass Spectrom.* **2003**, *14*, 825-833.

(29) Crowe, M. C.; Brodbelt, J. S. In *15th Sanibel Conference on Mass Spectrometry*: Sanibel, FL, 2003.

(30) Black, D. M.; Stephens, J. D.; Glish, G. L. In *Proceedings of the 51st ASMS Conference on Mass Spectrometry and Allied Topics*: Montreal, Canada, 2003.

(31) Keller, K. M.; Ostrander, C. M.; Fannin, S. T.; Brodbelt, J. S. In *Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics*: Chicago, IL, 2001.

(32) Reid, G. E.; Mitchell Wells, J.; Badman, E. R.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2003**, *222*, 243-258.

(33) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.

(34) McLuckey, S. A.; Habibi-Goudarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085-12095.

(35) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Processes* **1994**, *137*, 121-149.

(36) Barry, J. P.; Vouros, P.; Van Schepdael, A.; Law, S.-J. *J. Mass Spectrom.* **1995**, *30*, 993-1006.

(37) Wang, Z.; Wan, K. X.; Ramanathan, R.; Taylor, J. S.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 683-691.

(38) Wan, K. X.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 580-589.

(39) Wan, K. X.; Gross, J.; Hillenkamp, F.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 193-205.

(40) Premstaller, A.; Huber, C. G. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1053-1060.

(41) McLuckey, S. A.; Vaidyanathan, G.; Habibi-Goudarzi, S. *J. Mass Spectrom.* **1995**, *30*, 1222-1229.

(42) Sannes-Lowery, K. A.; Mack, D. P.; Hu, P.; Mei, H.-Y.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 90-95.

(43) Ni, J.; Mathews, M. A. A.; McCloskey, J. A. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 535-540.

(44) Wang, P.; Bartlett, M. G.; Martin, L. B. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 846-856.

(45) Weimann, A.; Iannitti-Tito, P.; Sheil, M. M. *Int. J. Mass Spectrom.* **2000**, *194*, 269-288.

(46) Vrkic, A. K.; O'Hair, R. A. J.; Foote, S.; Reid, G. E. *Int. J. Mass Spectrom.* **2000**, *194*, 145-164.

(47) Boschenok, J.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 144-149.

(48) Hakansson, K.; Hudgins, R. R.; Marshall, A. G.; O'Hair, R. A. J. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 23-41.

Chapter 5: Charge State-Dependent Fragmentation of Oligonucleotide/Metal Complexes

5.1 INTRODUCTION

Metal ions are critical to the *in vivo* structure and function of DNA and RNA. Mono- and divalent cations are thought to induce bending in DNA duplexes,[1,2] and are also known to stabilize triplex[3,4] and quadruplex[5,6] structures. Metal cations are also required for the proper folding and function of many forms of RNA, including most ribozymes.[7,8]. The interactions of DNA with several natural products[9] and synthetic drugs and drug candidates[10,11] are known to be metal-mediated.

Known as a useful tool for structural characterization and sequencing of nucleic acids,[12-14] mass spectrometry (MS) has also been employed to characterize a number of nucleic acid/metal ion interactions.[15] Several previous studies have focused on the interactions of the anticancer therapeutic cisplatin or its derivatives with DNA.[16-22] Relatively few studies have addressed binary interactions of nucleic acids and metal cations, however.[23-27] Most of these studies have focused primarily on complexes with the smaller alkali and alkaline earth metals (e.g. Na^{+}, Mg^{+2}), although a few complexes with a transition metals[25,26] or f-block elements[23,25] have also been investigated. The fragmentation behavior of these complexes has been characterized using parent species in relatively low charge state (often –2), and the charge state-dependence of the observed dissociation pathways has not been evaluated.

The present study explores the effect of metal complexation on the fragmentation patterns of oligodeoxynucleotides (ODN's) in several different charge states. A number

of different alkali, alkaline earth, and transition metals and several ten-residue sequences were employed to form 1:1 complexes, which were subjected to collision-activated dissociation (CAD) in a quadrupole ion trap mass spectrometer. For low charge state precursors, distinct changes were observed in the relative intensities of some product ions for free vs. metallated ODN's. For high charge state precursors, the metal complexes of some sequences exhibited unusual but highly diagnostic fragmentation pathways not commonly observed for free ODN's. The sequence and metal dependence of these results were evaluated to more fully expose the factors controlling these fragmentation processes.

5.2 EXPERIMENTAL

HPLC-purified oligonucleotides (ODN's) were obtained from TriLink Biotechnologies (San Diego, CA), Invitrogen (Carlsbad, CA), or IDT (Coralville, IA). When necessary ODN's were converted to the ammonium salt form by either centrifugal filtration or ethanol precipitation from ammonium acetate. Alkali, alkaline earth, and transition metal salts (bromides or chlorides) were obtained from Aldrich (St. Louis, MO) and were used as received. Samples containing equimolar amounts of oligonucleotide and metal salt at 20-40 μM in 1:3 methanol:water were infused at 3 μL/minute into the electrospray source of a ThermoQuest LCQ Duo quadrupole ion trap mass spectrometer operating in negative ion mode. In CAD experiments, activation voltages were applied at a level required to reduce the precursor ion to ~10% of its original intensity.

5.3 RESULTS AND DISCUSSION

Oligonucleotides are known to bind metal cations with high affinity. When equimolar amounts of metal salts were added to the ammonium form of a model oligonucleotide, the decamer GCGAATTCGC, ODN/metal complexes were readily observed for a wide variety of mono- and divalent metal cations, including K^+, Rb^+, Cs^+, Ca^{+2}, Sr^{+2}, Ba^{+2}, Mn^{+2}, Co^{+2}, Ni^{+2}, Ag^+, Cd^{+2}, Hg^{+2}, and Pb^{+2}. Several representative ESI spectra are shown in Figure 5.1. For all metals investigated, the 1:1 ODN:metal stoichiometry was observed in multiple charge states (typically –2 to –6), although higher stoichiometries were also observed for the late transition metals (Ni^{+2}, as shown in Fig. 5.1D; also Co^{+2} and Ag^+). The exact mode by which metal cations bind DNA remains unclear, but some evidence indicates that transition metals can bind DNA nucleobases, while alkali and alkaline earth metals prefer the backbone phosphodiester groups.[28] The relative intensities for 1:1 ODN:metal complexes in Fig. 5.1 could be interpreted as support for this notion: transition metals may produce higher stoichiometries and greater signal intensities because they can readily bind the bases and therefore need not compete with the ammonium counterions for phosphodiester binding sites. Alkali and alkaline earth cations, on the other hand, may compete less effectively for phosphodiester sites occupied by ammonium counterions, which are present in excess relative to the added metal salts.

To evaluate the effect of different metals on oligonucleotide fragmentation, CAD spectra were first acquired for the $[ODN - nH]^{-n}$ ions (n = -2 to -6) formed with GCGAATTCGC. CAD experiments were then performed for the analogous $[ODN - nH + M^{+m}]^{-(n-m)}$ ions (n – m = -2 to -6) formed by 1:1 complexes of GCGAATTCGC with a

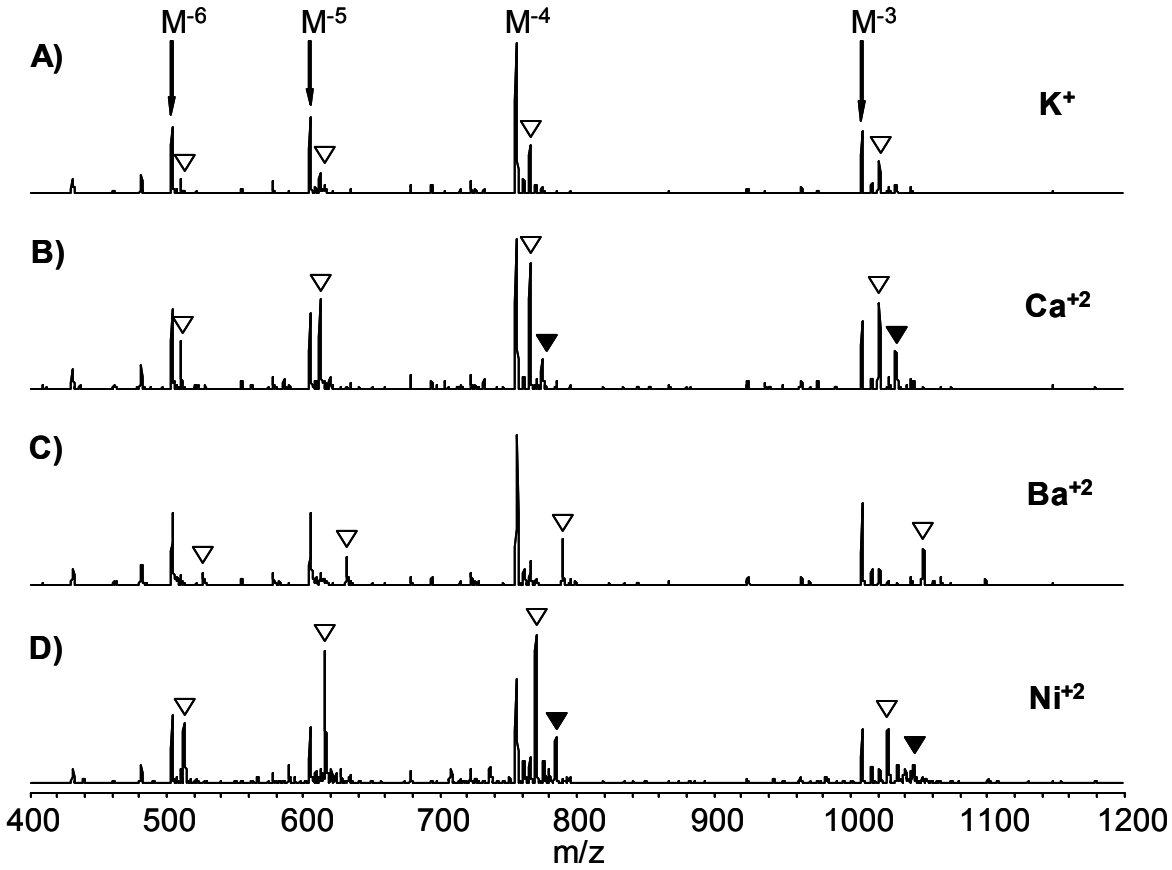


Figure5.1: ESI mass spectra for GCGAATTCGC in the presence of various metal salts. A) KCl, B) $CaCl_2$, C) $BaCl_2$, and D) $NiBr_2$. Complexes at the 1:1 stoichiometry are marked with open triangles (▽), and complexes at the 1:2 stoichiometry are marked with filled triangles (▼).

series of metal cations. Compared to the deprotonated precursors, some of the metal adducts demonstrated distinctly different fragmentation behavior in both low and high charge states. The metals most likely to produce unusual fragmentation were identified, and a representative metal (Ba^{+2}) was chosen for additional experiments with different oligonucleotides to investigate the sequence dependence of these dissociation processes.

5.3.1 Fragmentation of Low Charge State Precursors

Representative dissociation data for the –2 charge states of GCGAATTCGC and its 1:1 complex with Ba^{+2} are shown in Figure 5.2. For the deprotonated ODN, the most abundant ions in the spectrum represent loss of neutral nucleobases (detailed in Fig. 5.2A, inset), resulting in M-B ions that retain the intact phosphate backbone. Secondary dissociation of these ions produced an array of backbone cleavage products ("sequence ions") labeled w_n and a_n-B according to the McLuckey nomenclature.[29] The general mechanism shown in Scheme 5.1 has been proposed to explain the formation of these kinds of fragments.[30,31] The assortment of sequence ions visible in Fig. 5.2A indicates that cleavage has occurred at nearly every interresidue junction along the phosphate backbone, as indicated by the series of slash marks overlaid on the sequence shown in Fig. 5.2A. As originally described by Favre et al[32], the charges on these sequence ions convey the preferred locations of the negative charges. For example, the w ion series in Fig. 5.2A consisted of w_2^{-1}, w_5^{-1}, w_6^{-1}, w_8^{-2}, and w_9^{-2} species. The fact that w_5^{-2} and w_6^{-2} species were not observed suggests that in the activated parent species, the total charge was not distributed statistically along the phosphate backbone, but was instead confined near the 5'- and 3'-termini. Analysis of the a-B ion series in Fig. 5.2A supports the same conclusion, which is also consistent with results for different sequences in previous studies.[32]

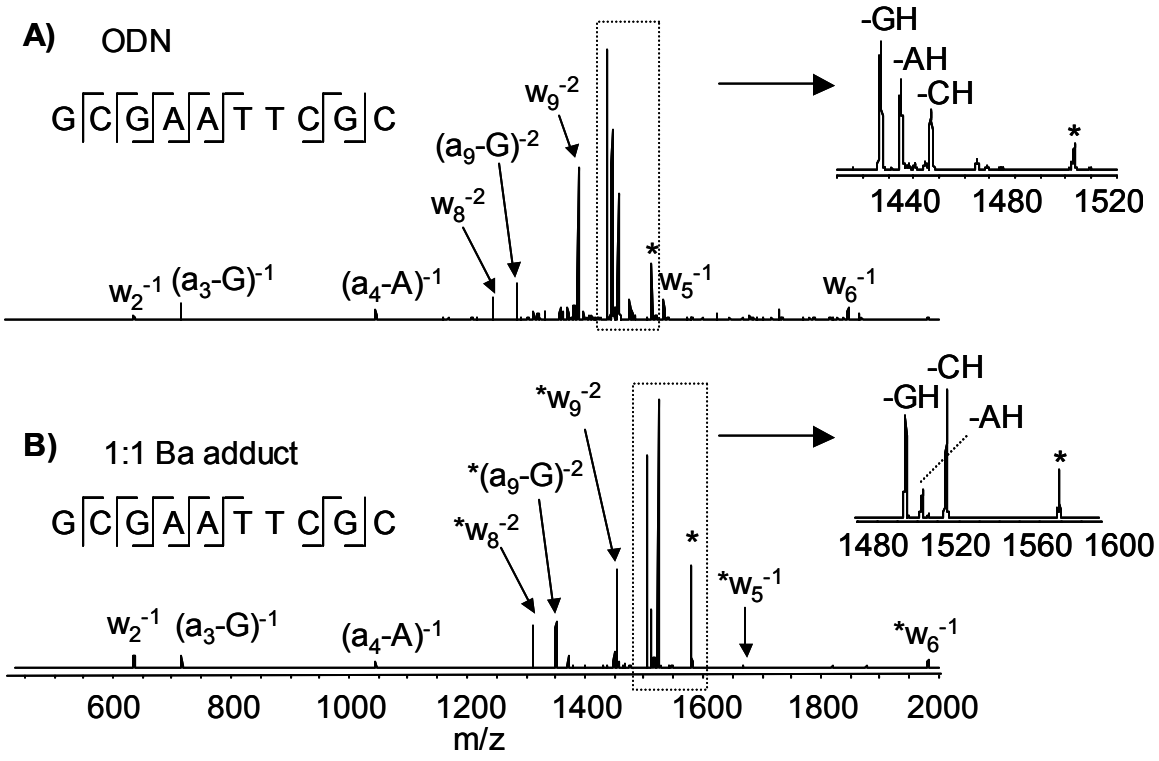
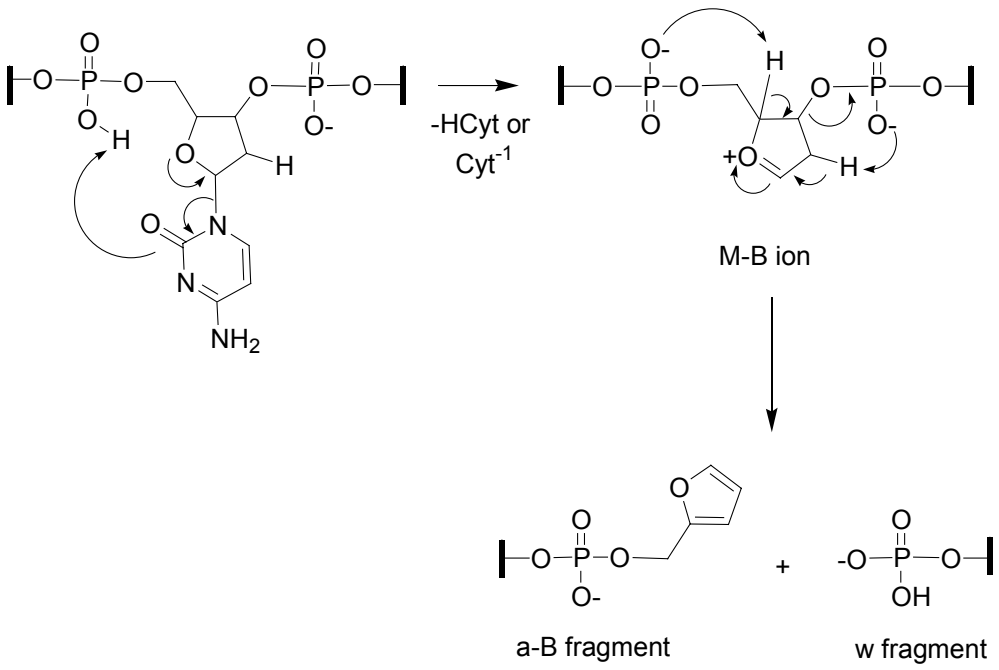


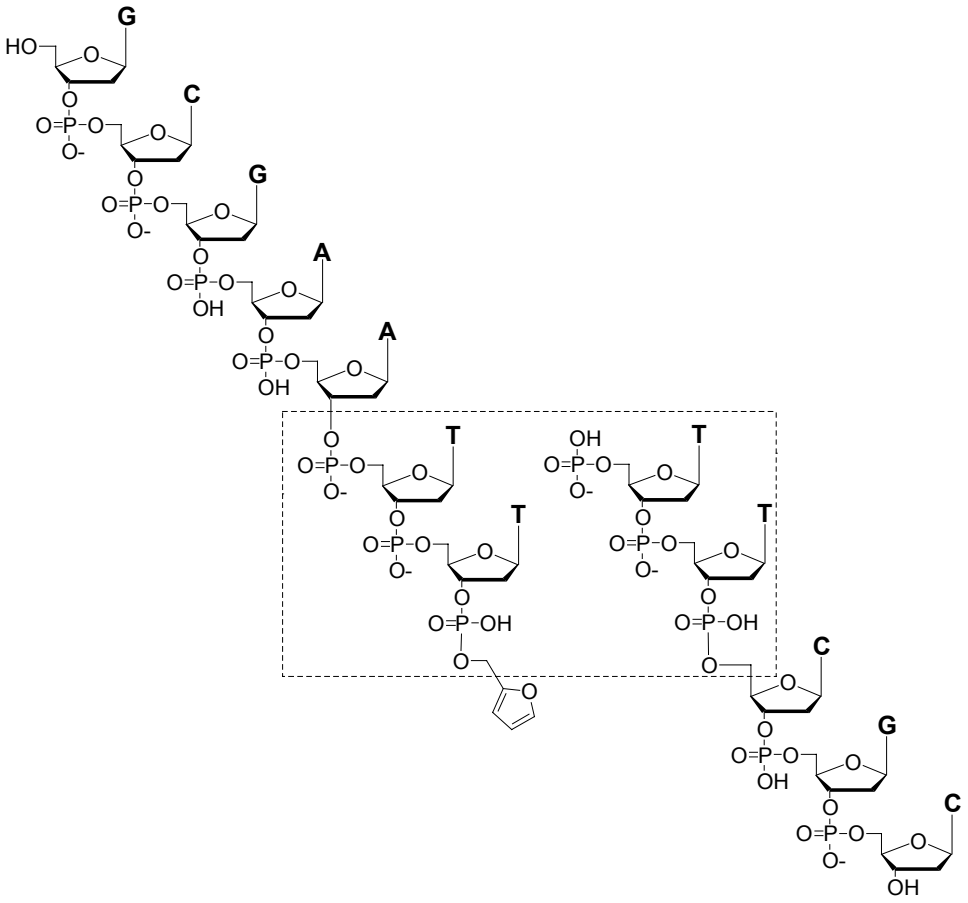
Figure 5.2: CAD spectra for the –2 charge state of GCGAATTCGC and its 1:1 complex with Ba^{+2}. A) ODN alone (0.75V/30 ms) and B) Ba^{+2} complex (0.88V/30 ms). Parent ions are marked *, and fragments identified with an asterisk and a standard sequence ion identifier (e.g. $*w_8^{-2}$) correspond to sequence ions that retain the metal ion.



Scheme 5.1: General mechanism for oligonucleotide base loss and backbone cleavage.

The CAD spectrum acquired for the 1:1 barium complex of the same decamer is shown in Fig. 5.2B. The m/z values for many of the resulting fragments indicated that the ODN-Ba^{+2} interaction survived collisional activation. These fragments include the M-B species (Fig. 5.2B inset) as well as a number of w and a-B ions, which are marked with asterisks to distinguish them from non-metallated w and a-B species. The combined array of both metallated and non-metallated sequence ions indicates that the Ba^{+2} complex produced the same overall sequence coverage as the free ODN (Fig. 5.2A), which is consistent with previous dissociation studies conducted with ODN/metal complexes.[23,24] In addition, the distribution of charge among the sequence ions suggests that one negative charge was localized near each end of the sequence, as described above for the free ODN. It is important to note, however, that metal cations provide partial charge balance for species of the general formula $[ODN - nH + M^{+m}]^{-(n-m)}$, and that the fragment ion charges can only be used to determine the position of the "excess" charges. Additional deprotonation sites counterbalanced by the metal cation must also exist, and these cannot be determined solely from the charge states of the sequence ions. The preferred localization of the metal cation itself can be estimated, however, by examining which sequence ions retain the metal.[24] In Fig. 5.2B, for example, the smallest w species retaining Ba^{+2} was the $*w_5^{-1}$ ion and the smallest a-B species retaining Ba^{+2} was the $*(a_8\text{-}C)^{-2}$ ion. The approximate location of the Ba^{+2} ion should correspond to the portion of the original sequence common to both these fragments (i.e. the pTpTp motif), as illustrated in Scheme 5.2. The fact that the metal resides near the middle of the sequence is consistent with results reported for series of sodium-adducted ODN's.[24]

In several respects (sequence coverage, the location of the negative charges, and the location of the metal cation), the data in Figure 5.2 correlate well with previous observations as noted above. The fragmentation behavior of the Ba^{+2} adduct was unusual



Scheme 5.2: Identification of the approximate binding site for Ba^{+2} on GCGAATTCGC.

in one respect, as illustrated by the insets in Fig. 5.2A and B, which are expansions of the base loss region of the CAD spectra. For the deprotonated ODN, the most intense M-B ion represented loss of neutral guanine, which is typical of mixed-base oligonucleotide sequences in low charge states.[33] For the Ba^{+2} adduct, however, loss of cytosine was the predominant process as indicated by the enhanced intensity of the M-CH ion. As indicated in Scheme 5.1, M-B species are intermediates that dissociate further to yield sequence ions. While loss of cytosine in and of itself is not atypical, loss of either adenine or guanine usually predominates, although the reasons for this are not well understood.[34] The enhanced loss of cytosine seen in Fig. 5.2A was found to depend on the identity of the metal complexed to the ODN, as illustrated in Fig. 5.3, which shows the base loss regions of the CAD mass spectra for ODN complexes containing different metal ions. Collisional activation of alkali or transition metal complexes produced different base loss patterns than the free ODN (Fig. 5.3A), as indicated by representative spectra for K^+ and Ni^{+2} (Fig. 5.3C and E). However, loss of cytosine was most strongly favored by complexation with the alkaline earth metals. The relative intensity of the M-CH ions were greatest for 1:1 adducts with Ba^{+2}, Sr^{+2} and Ca^{+2}; CAD spectra for Ba^{+2} and Ca^{+2} are shown in Fig. 5.3B and D, respectively.

Previous studies have shown that loss of cytosine is enhanced for fully metallated oligonucleotides, i.e. those where all phosphodiester counterions are metals.[27] To explain this enhancement the mechanism illustrated in Scheme 5.3 was proposed and supported with CAD results for deuterated precursor ions.[27] In this route, protons are transferred to the nucleobase from the 2' position of the deoxyribose ring, rather than the 5'-phosphate group as in the standard mechanism (Scheme 5.1). Subsequent base loss and cleavage at the 3' phosphate then generates the usual w and a-B sequence ions. Under these conditions, loss of cytosine and thymine typically predominates over loss of adenine and

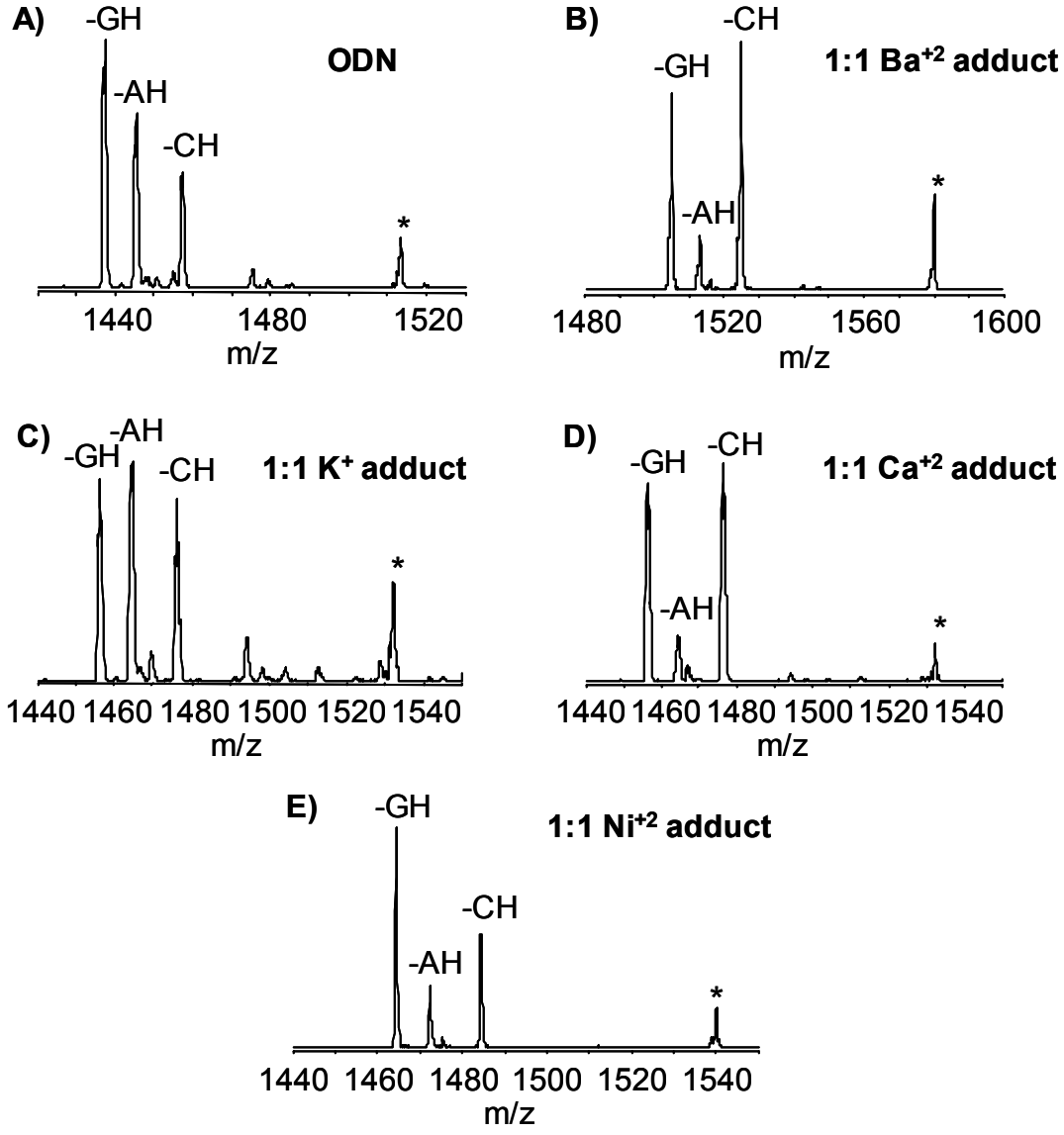
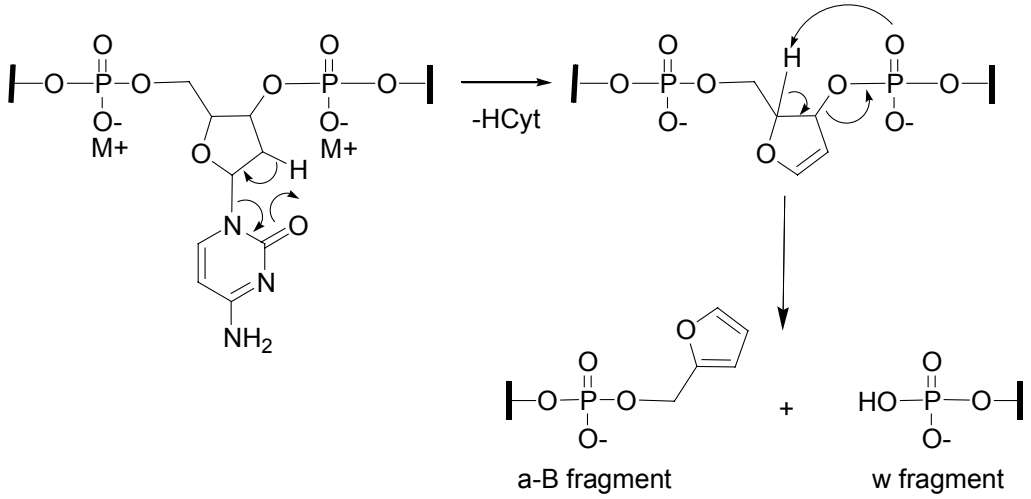


Figure 5.3: Base loss regions of CAD spectra for GCGAATTCGC and several 1:1 metal adducts in the –2 charge state. A) ODN alone (0.75V/30 ms), B) 1:1 adduct with Ba^{+2} (0.85V/30 ms), C) 1:1 adduct with K^{+} (0.95V/30 ms), D) 1:1 adduct with Ca^{+2} (0.9V/30 ms), and E) 1:1 adduct with Ni^{+2} (0.78V/30 ms). Parent ions are marked *.



Scheme 5.3: Proposed mechanism for base loss from ODN's bearing metal counterions.

guanine because the pyrimidines possess tautomeric forms that place electron-rich moieties in the proximity of the ribose ring. The mechanism depicted in Scheme 5.3 does not fully explain the dissociation patterns observed in Figs. 5.3B-E, however. This route predicts facile loss of thymine as well as cytosine, and essentially no loss of thymine was observed for any of the metal complexes investigated in our study. Furthermore, this mechanism was proposed to be operative at much higher metal:ODN stoichiometries, i.e. when all backbone phosphate protons are replaced with metal counterions. The parent ions in Figs. 5.3B and D contain only one metal cation, and therefore retain several acidic protons available for transfer to a nucleobase.

In an attempt to gain more insight into these unusual base losses, additional experiments were undertaken to determine what if any sequence dependence might exist. Fragmentation patterns were acquired for the –2 charge state of several other 10-residue oligonucleotides and with their 1:1 complexes with Ba^{+2}. To facilitate comparisons with the data in Fig. 5.2, these additional sequences all shared the base composition $A_2C_3G_3T_2$. CAD results obtained for the ODN's GAATTCGCGC, GCGCGAATTC, CGCTTAAGCG, and AAGCGCGCTT, were quite similar to those shown in Fig. 5.2: loss of guanine dominated for the $[M - 2H]^{-2}$ precursors, and loss of cytosine dominated for the $[M - 4H + Ba^{+2}]^{-2}$ precursors (data not shown). This uniform behavior suggests that the position of the cytosine residues along the sequence played little role in the observed fragmentation behavior. However, Ba^{+2} adduction did not significantly affect the base loss pattern for the sequence ACGTCTGCAG, as shown in Figure 5.4. In this case, the loss of guanine dominated the CAD spectra of both the free ODN (Fig. 5.4A) and the 1:1 barium adduct (Fig. 5.4B), and the loss of cytosine was comparable for both parent ions. It is not entirely clear why the CAD patterns for ACGTCTGCAG differ from the patterns observed for the other 10mers. However, for all five sequences for

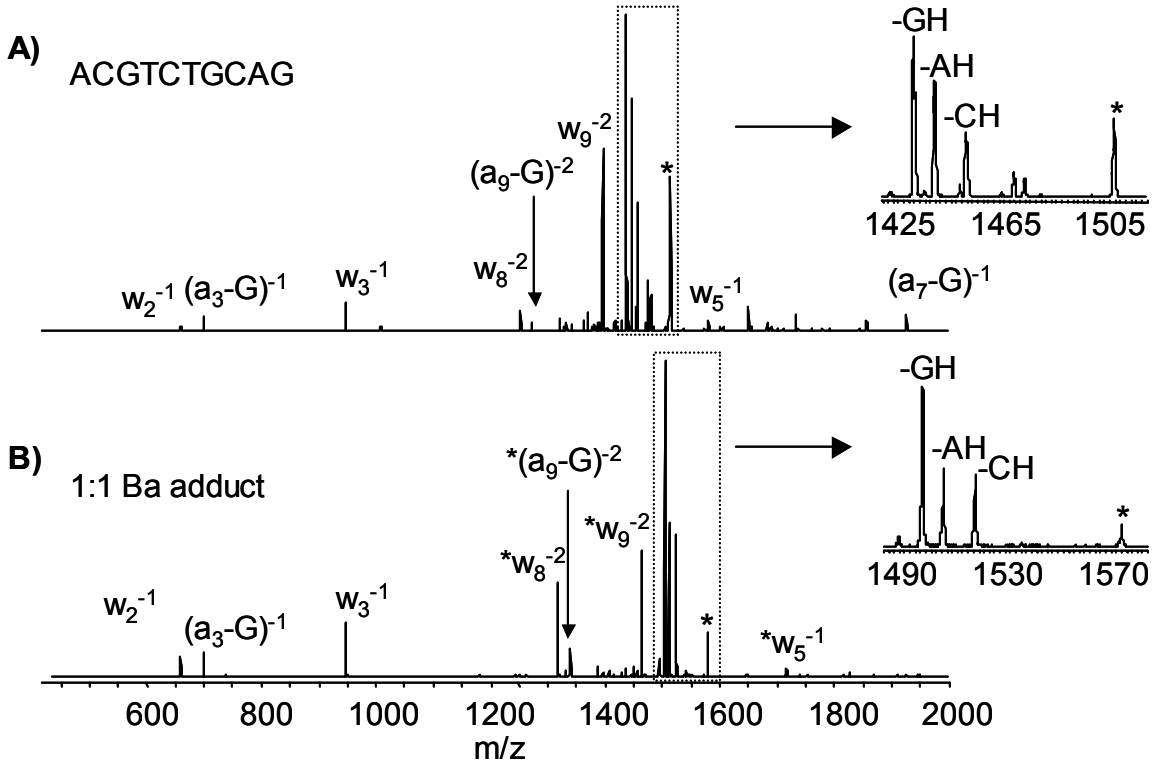


Figure 5.4: CAD spectra for the –2 charge state of ACGTCTGCAG and its 1:1 complex with Ba^{+2}. A) ODN alone (0.75V/30 ms) and B) Ba^{+2} complex (0.78V/30 ms). Parent ions are marked *, and fragments identified with an asterisk and a standard sequence ion identifier (e.g. $*w_8^{-2}$) correspond to sequence ions that retain the metal ion.

which the loss of cytosine was enhanced upon barium complexation, the residue at the 3'-terminus was the Watson-Crick complement of the residue at the 5'-terminus. This is not true of ACGTCTGCAG, the only sequence that did not exhibit a change in base loss pattern. This points to intramolecular base pairing, and therefore gas-phase conformation, as a significant factor: metal complexation may change the gas-phase conformation (and/or conformational dynamics) for some sequences, providing access to alternative mechanisms for proton transfer and subsequent base loss. Relatively little is known about the gas-phase conformation of oligonucleotides, although some results from preliminary ion mobility experiments have recently appeared.[35-39]

Fragmentation data for GCGAATTCGC and its 1:1 Ba^{+2} complex in the –3 charge state (not shown) produced $(M-BH)^{-3}$ species with intensity patterns similar to those visible in Fig. 5.2, which indicates that the factors controlling base loss processes for ODN's vs. their barium complexes were also operative at a higher charge state. CAD spectra acquired for the –4 charge states of GCGAATTCGC and its 1:1 Ba^{+2} complex are given in Figure 5.5A and B, respectively, and the base loss patterns in Figs. 5.5A and B are again different. For this charge state, however, the parent ions lose adenine as the major process, which is entirely consistent with previous literature reports for collisional activation of deprotonated oligonucleotides in intermediate charge states.[33,40,41] As was observed for the –2 charge state, the Ba^{+2} complex in the –4 charge state produced the same overall sequence coverage as the deprotonated ODN alone, and analysis of the metallated sequence ions indicated that the metal cation again localized preferentially at the pTpTp motif of the parent ion. Comparison of Figs. 5.2B and 5.5B illustrates that the M-B species became less dominant for the metal complex as charge state increased. This may mean that the difference between the threshold energy levels required for base loss and (secondary) backbone cleavage is smaller for metal complexes in higher vs. lower

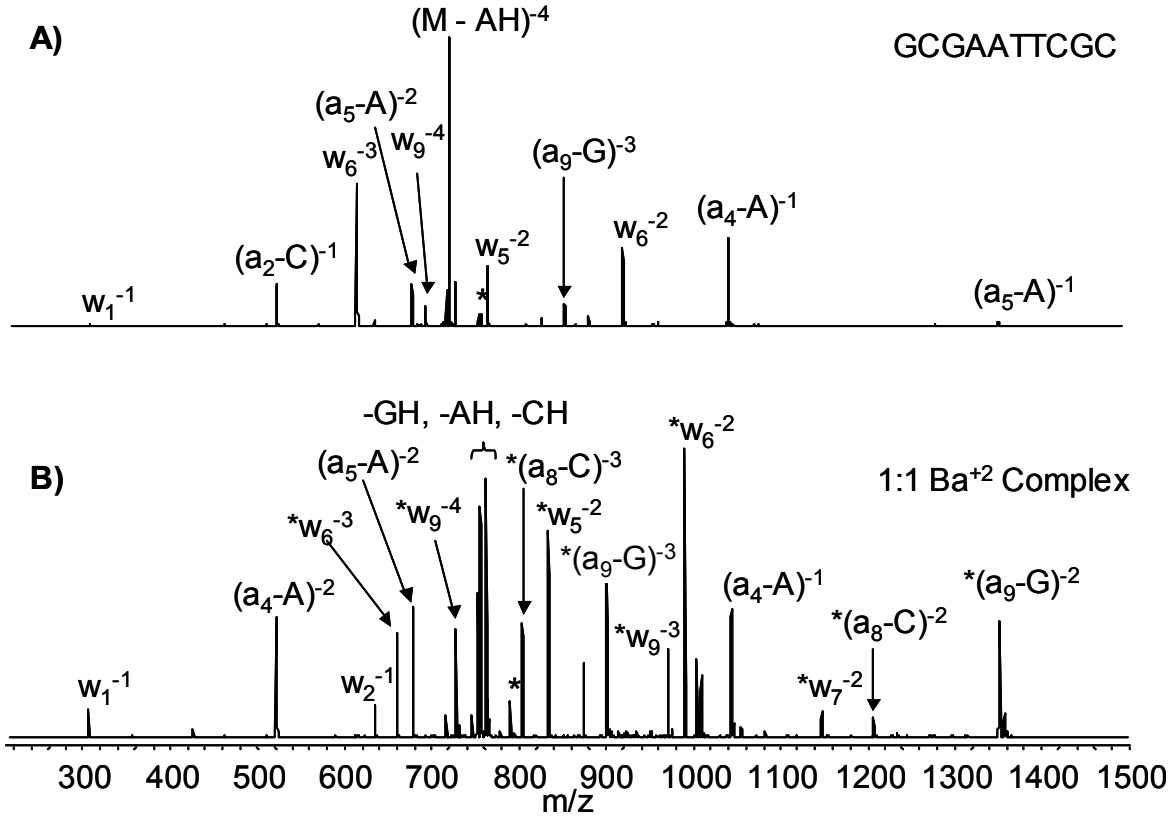


Figure 5.5: CAD spectra for the –4 charge state of GCGAATTCGC and its 1:1 complex with Ba^{+2}. A) ODN alone (0.75V/30 ms) and B) Ba^{+2} complex (0.80V/30 ms) . Parent ions are marked *, and fragments identified with an asterisk and a standard sequence ion identifier (e.g. $*w_6^{-3}$) correspond to sequence ions that retain the metal ion.

charge states. A similar difference in threshold energies may also explain the fact that sequence ions produced by the ODN:Ba^{+2} complex in Fig. 5.5B are more intense than the sequence ions visible in Fig. 5.5A. In addition, the intensity of sequence ions at higher m/z values than the parent ion was often greater in the spectrum of the Ba^{+2} complex than in the spectrum for the deprotonated ODN alone (ex. w_6^{-2} vs. $*w_6^{-2}$ in Figs. 5.5A and B).

5.3.2 Fragmentation of High Charge State Precursors

Fragmentation data for the –5 charge state of GCGAATTCGC and its 1:1 Ba^{+2} adduct are given in Figure 5.6A and B. The spectrum in Fig. 5.6B is remarkable in several respects: the M-B ions, which normally dominate CAD spectra for oligonucleotides acquired in quadrupole ion trap instruments, occur in unusually low signal intensity. Furthermore, the dominant product ions, at m/z 965.5 and 408.7, are consistent with the assignments $*a_6^{-2}$ and w_4^{-3}, respectively. The CAD spectra of oligonucleotides rarely contain a_n^{-n} ions as major species. The presence of the $*a_6^{-2}$ ion along with the complementary w_4^{-3} ion with similar signal intensity indicates that direct, site-specific backbone cleavage (no preceding base loss) has taken place. To confirm the assignment of m/z 965.5, an MS^3 experiment was performed, and the resulting spectrum (Figure 5.7) supports assignment of m/z 965.5 as the $*a_6^{-2}$ ion. Fragments assigned as $*a_5^{-1}$ and $*a_4^{-1}$ ions were observed along with the complementary [thymine monophosphate – $H_2O]^{-1}$ and $[pAT\text{-}H_2O]^{-1}$ ions, indicating that the $*a_6^{-2}$ ion also dissociated by direct backbone cleavage. Some products of cleavage preceded by base loss (likely by the mechanism shown in Scheme 5.2) are also visible (e.g. $*[a_6\text{-}T]^{-2}$) in Fig. 5.7.

The formation of the unusual $*a_6^{-2}$ species was found to be strongly metal dependent. The relative intensity of the $*a_6^{-2}/w_4^{-3}$ complementary pair decreased in the order Ba^{+2}, $Sr^{+2} > Ca^{+2}$, Pb^{+2}, $> K^+$, Rb^+, Cs^+, Mn^{+2}, Co^{+2}, Ni^{+2}, Cd^{+2}, Hg^{+2}.

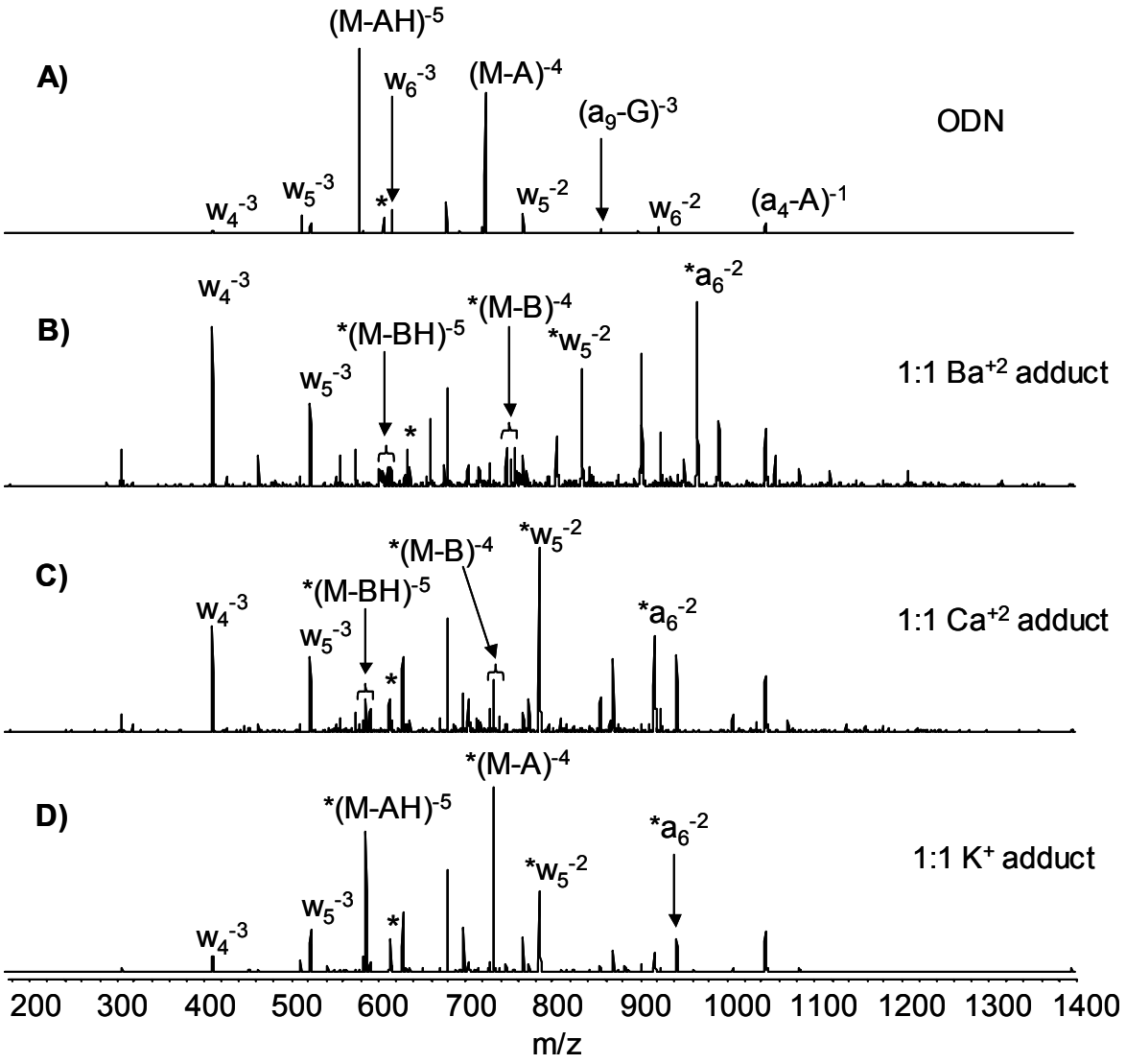


Figure 5.6: CAD spectra for the –5 charge state of GCGAATTCGC and its 1:1 complexes with several metal cations. A) ODN alone (0.75V/30 ms), B) Ba^{+2} complex (0.88V/30 ms), C) Ca^{+2} complex (0.88V/30 ms) , and D) K^{+} complex (0.90V/30 ms) . Parent ions are marked *, and fragments identified with an asterisk and a standard sequence ion identifier (e.g. $*w_5^{-2}$) correspond to sequence ions that retain the metal ion.

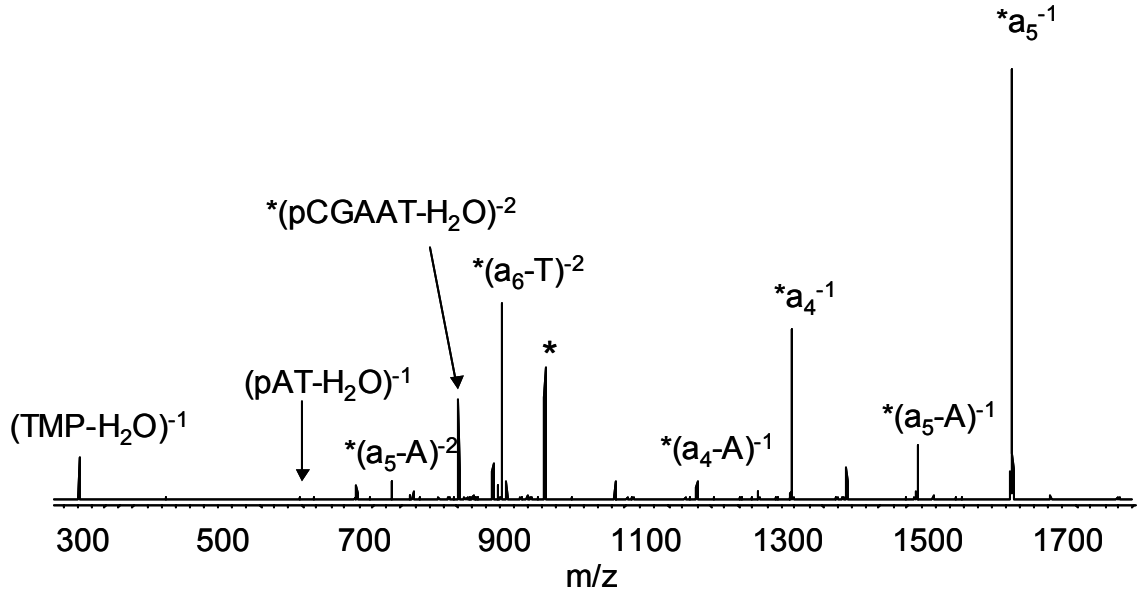


Figure 5.7: MS^3, $(GCGAATTCGC + Ba^{+2} - 7H)^{-5}$ → m/z 965.5 → ? (0.88mV/30 ms; 0.95V/30 ms). The parent ion is marked * and fragments identified with an asterisk and a standard sequence ion identifier (e.g. $*a_5^{-1}$) correspond to sequence ions that retain the metal ion.

Representative spectra acquired for complexes with Ca^{+2} and K^{+} are shown in Figure 5.6C and D, respectively. Comparing the CAD spectra for the complexes with Ba^{+2} (ionic radius 149 pm) and Ca^{+2} (114 pm) indicates that larger metal ions promote the formation of the $*a_6^{-2}$ product ion more effectively than smaller ions carrying the same ionic charge (Fig. 5.6B vs. C). Likewise, comparing data for Ba^{+2} and K^{+}, which have similar ionic radii (149 vs. 152 pm) but different net charges, indicates that greater ionic charge also facilitates formation of the $*a_6^{-2}$ product ion (Fig. 5.6B vs. D). Results for other metal ions suggests that other factors are probably also operative: for example, the complex with Pb^{+2} gave the $*a_6^{-2}$ ion in somewhat lower intensity than the complex with Sr^{+2}, which could be due to a difference in electronegativity (0.95 for Pb^{+2} vs. 1.87 for Sr^{+2}), orbital configuration (two vs. zero valence electrons), or both.

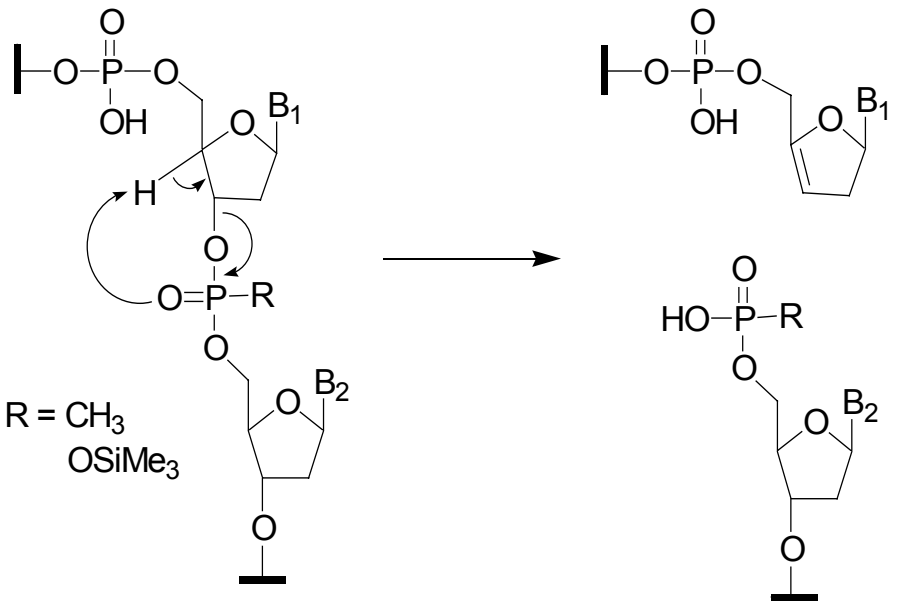
The formation of the $*a_6^{-2}$ species was also found to be strongly sequence dependent. Data for the relative intensity of the $*a_6^{-2}$ fragment for different ten-residue sequences is compiled in Table 5.1 and presented in a semi-quantitative fashion. The barium complexes of two unrelated sequences (ATGCTACGAG and GACTACAAGT) did not produce $*a_6^{-2}$ ions, showing that the direct backbone cleavage process is not a general phenomenon. The first five entries in Table 5.1, taken as a collective, indicate that the middle of the sequence greatly influences the direct backbone cleavage process. In a particularly striking result, replacing the central "AATT" motif in the original sequence with the alternating sequence "ATAT" completely prevented the formation of $*a_6^{-2}$ ion. The sequence GCAAATTTGC, which incorporates additional A and T residues, produced a strong $*a_6^{-2}$ ion, but the alternating version of this sequence (GCATATATGC) also gave the $*a_6^{-2}$ product with significant (albeit somewhat lower) intensity. The 10mer GCGCATGCGC, which contains fewer A/T residues, also produced the $*a_6^{-2}$ ion with substantial signal intensity. Taken together these results

Sequence	Relative Intensity, *a
GCGAATTCGC	+++
GCGATATCGC	-
GCAAATTTGC	+++
GCATATATCG	++
GCGCATGCGC	++
GCGAGTTCGC	++
GCGAACTCGC	+
GCGCGAATTC	+
GAATTCGCGC	+
CGCTTAAGCG	+
ATGCTACGAG	-
GACTACAAGT	-

Table 5.1: Sequence dependence of *a ion formation upon CAD of 1:1 Ba^{+2} complexes (-5 charge states).

indicate that the AT motif at positions 5 and 6 is critical, and that sequential A and T residues around these positions is supportive. Point substitutions at these sites, as in the sequences GCGAGTTCGC and GCGAACTCGC, show that the thymine residue at position 6 is more critical to the formation of the $*a_6^{-2}$ ion than the adenine at position 5. Three sequence variants (GCGCGAATTC, GAATTCGCGC, and CGCTTAAGCG) all produced metallated a_n^{-m} ions in only low intensity, which indicates that both the position and directionality of the AT motif are significant. This may mean that the proximity of the AT motif to the metal center may be critical, because for all of the decamers listed in Table 5.1, the distribution of the metal cation among the observed fragment ions indicates that the Ba^{+2} ions bound near the middle of the sequence (i.e. near positions 5 and 6).

Direct cleavage of the phosphate backbone of deprotonated oligonucleotides to yield a_n^{-m} ions has been observed in certain circumstances. Favre et al observed a number of a_n^{-m} ions upon collisional activation of deprotonated polythymine, a result that was ascribed to the low proton affinity of thymine.[32] ODN's with strictly uncharged backbone sites (either methylphosphonates[42] or TMS-derivatized phosphates[43]) have also been shown to undergo direct backbone cleavage at a site adjacent to the neutral linkage. One mechanism proposed to explain such cases[42], shown in Scheme 5.4, involves proton transfer from the 4' position of ribose to the neutral phosphodiester linkage followed by cleavage of the 3' ribose-linker bond. Presumably, the mechanism in Scheme 5.4 could also operate in cases where a standard but *neutral* phosphodiester linkage is present, i.e. when the R group in Scheme 5.4 = OH. Neutral backbone sites must be present in the ODN/Ba^{+2} complexes under investigation: the decamers contain a total of nine acidic (i.e. phosphodiester) hydrogens, so $[ODN - nH + Ba^{+2}]^{(n-2)}$ complexes carrying a net charge of –5 would retain two acidic protons on the phosphate backbone. These considerations lead to the hypothesis that for certain oligonucleotide/metal complexes,

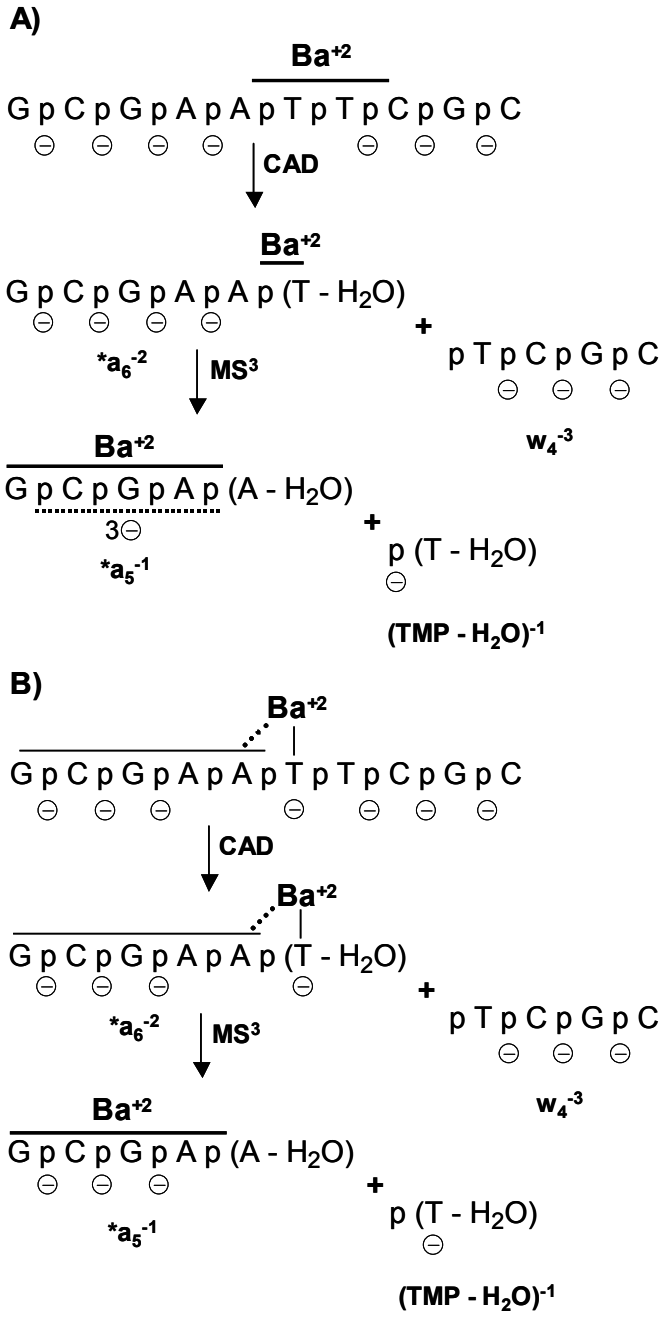


Scheme 4: Proposed mechanism for direct cleavage of the phosphate backbone to yield a_n^{-m} ions.

metal coordination may sequester acidic protons at certain positions along the phosphate backbone.

This concept is illustrated in Scheme 5.5A for the [d(GCGAATTCGC) – 7H + Ba^{+2}]$^{-5}$ ion. In accordance with CAD data for this complex (Fig. 5.6B), the Ba^{+2} cation is shown to be bound somewhere along the internal pTpTp subsequence. The negative charges are assumed to reside only on phosphodiester groups and are assigned to phosphodiesters 1-4 and 7-9 (relative to the 5'-terminus). This leaves neutral sites at phosphodiesters 5 and 6, which could lead to direct backbone cleavage by the mechanism in Scheme 5.4 to produce the $*a_6^{-2}$ ion along with its complement, the w_4^{-3} ion. The presence of a neutral phosphodiester at position 5 in the $*a_6^{-2}$ ion could then yield an a_5 product on MS^3, which was indeed observed (Fig. 5.7). Experimentally, however, the metallated a_5 species seen in Fig. 5.7 carries a net charge of –1, and the complementary (TMP – H_2O) ion (which contains the critical neutral phosphodiester) also carried a net charge of –1. This is difficult to reconcile by the route shown in Scheme 5.5A, where phosphodiester 5 is assigned as a neutral site, unless extensive proton transfer from other sites takes place.

An alternative route is presented in Scheme 5.5B. Here the possibility of deprotonation at both phosphodiester *and* nucleobase moieties is considered. (In general the nucleobases are far less acidic than the phosphodiester groups, but both guanosine and thymidine can be deprotonated in solution at high pH (pKa = 9.2 and 9.8, respectively).[44] In the gas phase, internal Coulombic repulsion in highly charged oligonucleotide anions could increase the likelihood of deprotonation of guanine and/or thymine residues.) If it is assumed that the critical thymine residue at position 6 is deprotonated and a specific interaction between this residue and the metal cation is maintained (allowing the possibility of additional interactions along the GCGAA motif),



Scheme 5.5: Possible route to *a ions from the -5 charge state of (GCGAATTCGC + Ba^{+2} - 7H)$^{-5}$. A) Negative charges assigned exclusively to phosphodiester groups and B) negative charges allowed on both phosphodiester groups and nucleobase moieties.

then a neutral site can be assigned to phosphodiester 4, as shown. Like the scenario in Scheme 5.5A, this permits direct cleavage of the parent species to produce an $*a_6^{-2}$ product ion according to the route in Scheme 5.4. However, Scheme 5.5B better reconciles the MS^3 data obtained for the $*a_6^{-2}$ ion in several respects. First, charge balance in the conversion of $*a_6^{-2}$ to $*a_5^{-1}$ + $(TMP - H_2O)^{-1}$ by Scheme 5.4 is more readily explained if phosphodiester 4 is neutral. Second, the presence of a formally charged thymine residue allows the possibility of anionic base loss, which could explain the large difference in relative intensities of the $*a_5^{-1}$ and $(TMP - H_2O)^{-1}$ ions in Fig. 5.7. (As complementary species, these ions would be expected to form with similar relative intensities. A $(T\text{-}H)^{-1}$ ion, if formed, would be seen at m/z 110, which is below the low mass cutoff of the spectrum in Fig. 5.7). Third, the presence of the $*a_4^{-1}$ product ion in Fig. 5.7 could form by the route in Scheme 5.4 if phosphodiester 4 is neutral. A reasonable route to this product ion is less readily conceived from Scheme 5.5A.

In sum, Scheme 5.5B presents a picture for the formation of metallated a_n^{-m} fragments from $[d(GCGAATTCGC) - 7H + Ba^{+2}]^{-5}$ ion that is generally consistent with what is known of the factors influencing the direct cleavage process. The low proton count agrees with the observed charge state dependence. The notion that a (deprotonated) thymine residue at position 6 might be involved is generally consistent with the sequence dependence indicated in Table 5.1. It should be acknowledged, however, that Scheme 5.5 rationalizes only the major fragmentation pathways observed for the $[d(GCGAATTCGC) - 7H + Ba^{+2}]^{-5}$ ion (and its $*a_6^{-1}$ product ion). Other, less significant pathways are also observed in both Fig. 5.3B and Fig. 5.7, implying that multiple populations with different deprotonation and/or metal complexation sites may well exist and interchange.

The possibility of multiple populations is also highlighted by CAD data obtained for 1:1 metal complexes in the –6 charge state. The Ba^{+2} complexes of only a few of the decamers listed in Table 5.1 (GCGAATTCGC, GCAAATTTGC and GCGAGTTCGC) produced detectable $*a_n^{-m}$ ions. A representative spectrum for the barium complex of GCGAATTCGC is shown in Figure 5.8, and illustrates that for this precursor charge state the metallated a_n^{-m} species formed in lower relative intensity but greater variety. If negative charges occur exclusively at phosphodiester moieties, then ions of the formula $[ODN - nH + Ba^{+2}]^{-6}$ would retain only one neutral backbone site. If this neutral site drives direct backbone cleavage, as Scheme 5.4 hypothesizes, the distribution of different $*a_n^{-m}$ ions in Fig. 5.8 indicates that this "last" neutral site delocalizes over several positions along the backbone.

5.4 CONCLUSIONS

For low to intermediate charge states, metal adduction altered the relative intensity of M-B species produced for several ten-residue ODN's, promoting loss of cytosine rather than loss of guanine. The relative intensities of sequence ions were largely unaffected. This behavior was most prevalent for alkaline earth cations Ca^{+2}, Sr^{+2}, and Ba^{+2}, and for isomeric ODN's with complementary residues at the 5'- and 3'-termini. This suggests that metal complexation may change the gas-phase conformation for some sequences, possibly providing access to alternative mechanisms for proton transfer to and subsequent loss of the nucleobase moieties.

In higher charge states, CAD of some metal adducts of GCGAATTCGC produced intense fragment ions corresponding to metallated a_6^{-2} ions. These species result from direct cleavage of the phosphate backbone (with no prior loss of nucleobase) and are not commonly observed for deprotonated ODN's. The formation of these ions from metal

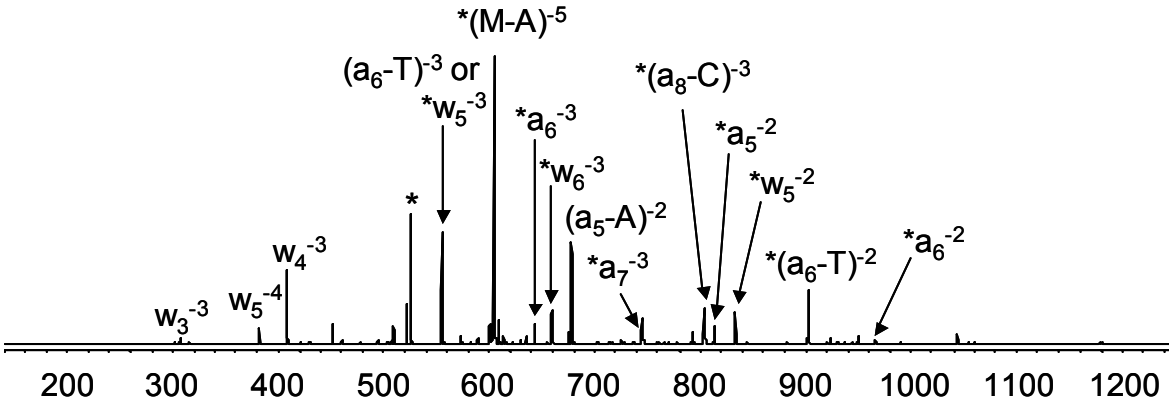


Figure 5.8: CAD spectra for the 1:1 adduct of GCGAATTCGC and Ba^{+2} in the -6 charge state (0.90V/30 ms). The parent ion is marked *, and fragments identified with an asterisk and a standard sequence ion identifier (e.g. $*w_5^{-3}$) correspond to sequence ions that retain the metal ion.

adducted ODN's is highly dependent on the identity of the metal, being most favored by large divalent cations with empty valence shells such as Ba^{+2} and Sr^{+2}. This behavior was also strongly dependent on sequence, being most prevalent for sequences with a thymine residue at position 6. Literature precedent exists for the formation of a_n^{-m} ions from sequences in which covalent modification generates one or more neutral sites along the phosphate backbone. ODN/metal adducts in high charge states possess only a few acidic protons, and the juxtaposition of these neutral phosphate groups near thymine residues and the bound Ba^{+2} ion may direct formation of the metallated a_6^{-2} species.

5.5 REFERENCES

(1) Hud, N. V.; Polak, M. *Curr. Opin. Struct. Biol.* **2001**, *11*, 293-301.

(2) McFail-Isom, L.; Sines, C. C.; Williams, L. D. *Curr. Opin. Struct. Biol.* **1999**, *9*, 298-304.

(3) Frank-Kamenetskii, M. D.; Mirkin, S. M. *Annu. Rev. Biochem.* **1995**, *64*, 65-95.

(4) Gilbert, D. E.; Feigon, J. *Curr. Opin. Struct. Biol.* **1999**, *9*, 305-314.

(5) Chowdhury, S.; Bansal, M. *J. Biomol. Struct. Dyn.* **2000**, *18*, 11-28.

(6) Hardin, C. C.; Perry, A. G.; White, K. *Biopolymers* **2000**, *56*, 147-194.

(7) Draper, D. E.; Shiman, R. *J. Mol. Biol.* **2000**, *302*, 79-91.

(8) Draper, D. E. *RNA* **2004**, *10*, 335-343.

(9) Chakrabarti, S.; Bhattacharyya, D.; Dasgupta, D. *Biopolymers* **2000**, *56*, 85-95.

(10) Fan, J. Y.; Sun, D.; Yu, H.; Kerwin, S. M.; Hurley, L. H. *J. Med. Chem.* **1995**, *38*, 408-424.

(11) Zeng, Q.; Kwok, Y.; Kerwin, S. M.; Mangold, G.; Hurley, L. H. *J. Med. Chem.* **1998**, *41*, 4273-4278.

(12) Ni, J.; Pomerantz, S. C.; Rozenski, J.; Zhang, Y.; McCloskey, J. A. *Anal. Chem.* **1996**, *68*, 1989-1999.

(13) Nordhoff, E.; Kirpekar, F.; Roepstorff, P. *Mass Spectrom. Rev.* **1996**, *15*, 69-138.

(14) Limbach, P. A. *Mass Spectrom. Rev.* **1997**, *15*, 297-336.

(15) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.

(16) Gonnet, F.; Kocher, F.; Blais, J.; Bolbach, G.; Chottard, J.; Tabet, J. C. *Inorg. Chem.* **1996**, *35*, 1653-1658.

(17) Lowe, G.; McCloskey, J. A.; Ni, J.; Vilaivan, T. *Bioinorg. Med. Chem.* **1996**, *4*, 1007-1013.

(18) Troujman, H.; Chottard, J. C. *Anal. Biochem.* **1997**, *252*, 177-185.

(19) Xu, N.; Pasa-Tolic, L.; Smith, R. D.; Ni, S.; Thrall, B. D. *Anal. Biochem.* **1999**, *272*, 26-33.

(20) Iannitti-Tito, P.; Weimann, A.; Wickham, G.; Sheil, M. M. *Analyst* **2000**, *125*, 627-634.

(21) Kloster, M. B. G.; Hannis, J. C.; Muddiman, D. C.; Farrell, N. *Biochemistry* **1999**, *38*, 14731-14737.

(22) Carte, N.; Legendre, F.; Leize, E.; Potier, N.; Reeder, F.; Chottard, J. C.; Dorsselaer, A. *Anal. Biochem.* **2000**, *284*, 77-86.

(23) Wu, Q.; Cheng, X.; Hofstadler, S. A.; Smith, R. D. *J. Mass Spectrom.* **1996**, *31*, 669-675.

(24) Favre, A.; Gonnet, F.; Tabet, J.-C. *Int. J. Mass Spectrom.* **1999**, *190*, 303-312.

(25) Hettich, R. L. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 941-949.

(26) Hettich, R. L. *Int. J. Mass Spectrom.* **2001**, *204*, 55-75.

(27) Wang, Y.; Taylor, J. S.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 550-556.

(28) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I. J. *Physical Chemistry of Nucleic Acids*; Harper and Row: New York, 1974.

(29) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.

(30) Wang, Z.; Wan, K. X.; Ramanathan, R.; Taylor, J. S.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 683-691.

(31) Wan, K. X.; Gross, J.; Hillenkamp, F.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 193-205.

(32) Favre, A.; Gonnet, F.; Tabet, J.-C. *Eur. J. Mass Spectrom.* **2000**, *6*, 389-396.

(33) Luo, H.; Lipton, M. S.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 195-199.

(34) Daneshfar, R.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 55-64.

(35) Hoaglund, C. S.; Liu, Y.; Ellington, A. D.; Pagel, M.; Clemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 9051-9052.

(36) Gidden, J.; Bushnell, J. E.; Bowers, M. T. *J. Am. Chem. Soc.* **2001**, *123*, 5610-5611.

(37) Gidden, J.; Bowers, M. T. *Eur. Phys. J. D* **2002**, *20*, 409-419.

(38) Gidden, J.; Bowers, M. T. *J. Phys. Chem. B* **2003**, *107*, 12829-12837.

(39) Gidden, J.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 161-170.

(40) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

(41) Keller, K. M.; Brodbelt, J. S. *Anal. Biochem.*. **2004**, *326*, 200-210.

(42) Bartlett, M. G.; McCloskey, J. A.; Manalili, S.; Griffey, R. H. *J. Mass Spectrom.* **1996**, *31*, 1277-1283.

(43) O'Hair, R. A. J.; McLuckey, S. A. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *162*, 183-202.

(44) Clauwaert, J.; Stockx, J. *Z. Naturforsch. B* **1968**, *23*, 25-30.

Chapter 6: Influence of Initial Charge State on Fragmentation Patterns for Non-Covalent Drug/DNA Duplex Interactions

6.1 INTRODUCTION

A hallmark of electrospray ionization mass spectrometry (ESI-MS) is its ability to ionize non-covalent complexes without disrupting fragile intermolecular interactions.[1-5] Because of this property, ESI-MS has been used to investigate a number of non-covalent interactions known or thought to be important in the function of DNA and RNA, including DNA duplexes,[6-11] triplexes,[12] and quadruplexes,[12,13], as well as drug/DNA[7,8,14-21] or drug/RNA complexes.[22-26] In such studies ESI-MS has been used to assess binding stoichiometry, metal ion requirements for binding, and relative binding affinity, and has also been used to estimate solution binding constants.

A second important characteristic of ESI is that it tends to produce multiply charged ions for analytes possessing multiple protonation or deprotonation sites. Fragmentation patterns of proteins[27-30] and single-stranded oligonucleotides[31,32] are known to vary with the charge state of the parent ion. Previous studies[10,14] have shown that the collisional activation energy required to dissociate DNA duplexes varies inversely with charge state. In terms of fragmentation pathways, it has been shown that under fast heating conditions (e.g. in quadrupole/TOF instrumentation), non-covalent strand separation predominates for all but the lowest charge states, which can also undergo covalent bond cleavage.[11] For a given precursor, covalent cleavage pathways become more significant in spectra acquired on quadrupole ion trap (QIT) instrumentation, which reflect slow heating conditions.[11] Both covalent and non-covalent

dissociation have been reported for various DNA duplex/drug complexes in the QIT, and general correlations have been drawn between the predominant fragmentation pathway and the mode of drug binding to the duplex.[8] No information has been reported on how charge state affects the fragmentation patterns of DNA duplex/drug complexes, however. We have therefore investigated this issue using collision-activated dissociation (CAD) in a quadrupole ion trap mass spectrometer (QIT-MS). Several self-complementary 10-residue oligonucleotides were employed, in addition to a number of DNA-interactive drugs, including the intercalators daunomycin and nogalamycin, and the minor groove binding agents distamycin, netropsin, 4,6-diamidino-2-phenylindole, and Hoechst 33342, (structures shown in Figure 6.1). In general, the dissociation pathways exhibited by both the duplexes and the drug/duplex complexes were found to be markedly sensitive to initial charge state. Results for the drug/duplex complexes were interpreted in terms of the known binding mode of the drug and the observed dependence of fragmentation on activation time. Data for the quinolone antibiotic norfloxacin, which interacts with DNA in a fashion that is poorly understood, was also acquired and evaluated.

6.2 EXPERIMENTAL

All DNA-interactive drugs were purchased from Sigma (St. Louis, MO) and used as received. The self-complementary oligodeoxynucleotides (ODNs) 5'-GCGCATGCGC-3', 5'-GCGAATTCGC-3', and 5'-GCAAATTTGC-3' were purchased from TriLink Biotechnologies (San Diego, CA) as HPLC-purified ammonium salts. To form DNA duplexes, stock solutions containing ODN (1 mM) and ammonium acetate (1 M) were annealed at 90°C for ten minutes and slowly cooled to room temperature over twelve hours. Equimolar solutions containing 25 μM each of DNA duplex, drug, and

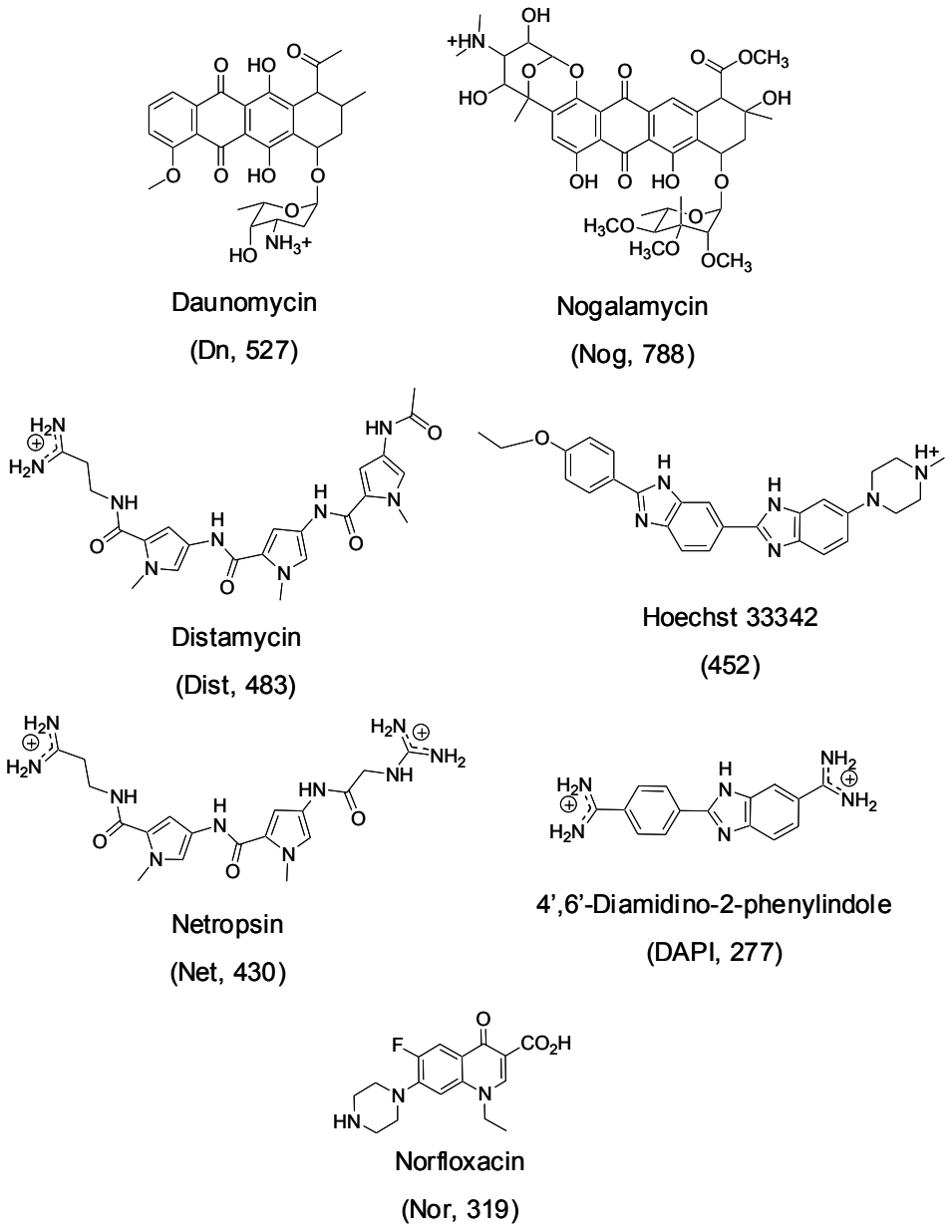


Figure 6.1. Structures of DNA-interactive drugs used in this study. Compound abbreviations and molecular weights in Da are given in parenthesis.

divalent metal salt (where applicable) were prepared in 25% methanol containing 50 mM ammonium acetate at a nominal pH of 6.5. The solutions were allowed to equilibrate at room temperature prior to analysis. Electrospray ionization mass spectra were collected on a ThermoQuest LCQ Duo ion trap mass spectrometer equipped with an electrospray ionization source operating in the negative ionization mode. The analyte solution was introduced into the instrument by direct infusion at a flow rate of 3 μL/min. A heated capillary temperature of 90°C was used to ensure adequate desolvation while minimizing in-source dissociation of the non-covalent complexes. In CAD experiments, activation voltages were applied at a level required to reduce the selected parent complex to ~10% of its original intensity. The instrument's default activation time of 30 ms was used except where noted. Since activation time and activation voltage vary inversely, greater activation voltages were required when activation times were decreased, and vice versa.

6.3 RESULTS AND DISCUSSION

Our objective was to unravel some of the parameters that influence the dissociation pathways of DNA/drug complexes in order to determine the extent to which binding modes are reflected in the CAD mass spectra. First, the CAD patterns of the duplexes alone were evaluated as a function of the rate of energy deposition (a function of the activation time and applied collisional activation voltage) in order to allow benchmark comparisons to the dissociation behavior of the duplex/drug complexes. Then the fragmentation patterns of a variety of duplex/drug complexes were studied to allow systematic investigation of the impact of the drug, the oligonucleotide sequence, the collisional activation time, and the stoichiometry of the duplex/drug complexes on the observed dissociation pathways.

6.3.1 Dissociation of DNA Duplexes

Gas-phase DNA duplexes are known to dissociate in the gas phase by one of two pathways: non-covalent cleavage (i.e. strand separation or "unzipping") to yield two single-stranded species, or covalent cleavage to yield either fragments of single strands (Type I products) or partial duplexes from which either base loss or backbone fragmentation of one strand has occurred (Type II products).[11] Previous studies have shown that the tendency for covalent cleavage increases with the number of interstrand hydrogen bonds, i.e. for longer vs. shorter duplexes and GC- vs. AT-rich sequences.[6,8,9,11] Low energy collisional activation also promotes covalent cleavage, so instrument platforms that provide slow heating conditions are more likely to yield covalent vs. non-covalent dissociation products.[11]

To study the effect of charge state on duplex fragmentation, both $[ds]^{-5}$ and $[ds]^{-4}$ species were generated by electrospray ionization of $d(GCGAATTCGC)_2$ (Fig. 6.2A). (Although the ion at m/z 1513 cannot be unambiguously assigned as $[ds]^{-4}$ or $[ss]^{-2}$ based solely on its mass-to-charge ratio, examination of the spacing of the associated sodium adduct peaks suggests that $[ds]^{-4}$ is the major species.) Collisional activation of the $[ds]^{-5}$ ion using a CAD time of 30 ms produced only non-covalent strand separation (Fig. 6.2B), a result that was found to be insensitive to activation time (data not shown). In contrast, CAD of the $[ds]^{-4}$ ion using a 30 ms activation period (Fig. 6.2D) gave several Type II products (e.g. $[ds - GH]^{-4}$, $[ds - a_4]^{-3}$, etc.), indicating that covalent cleavage of the duplex was a major process under these conditions. This result was found to be strongly dependent on the activation time: a shorter activation period and correspondingly greater collisional activation voltage (3 ms, 1.3 V, Fig. 6.2C) gave products of single strand dissociation (Type I products), while a longer activation time and lower activation

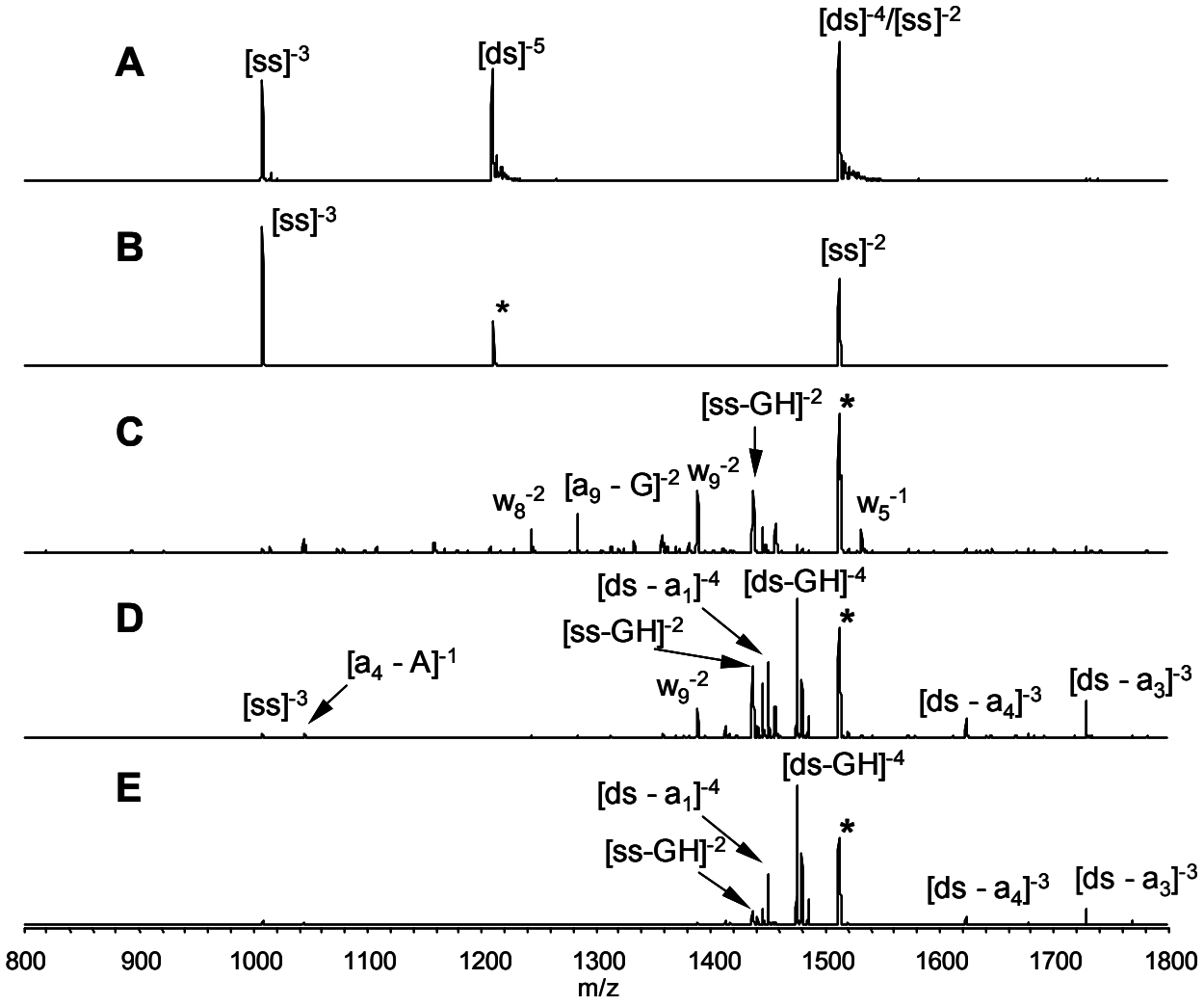


Figure 6.2: ESI and CAD product ion spectra for $d(GCGAATTCGC)_2$: (A) full scan mass spectrum, (B) CAD of $[ds]^{-5}$, 0.5V/30 ms activation, (C) CAD of $[ds]^{-4}$, 1.3V/3 ms activation (D) CAD of $[ds]^{-4}$, 0.70V/30 ms activation, and (E) CAD of $[ds]^{-4}$, 0.50V/300 ms activation. Parent ions are marked *.

voltage (300 ms, 0.50 V, Fig. 6.2E) gave mostly Type II fragments of duplex dissociation. The dominance of the $[ds - GH]^{-4}$ ion at m/z 1475 in Fig. 6.2E confirms that the parent ion population at m/z 1513 consisted almost exclusively of the $[ds]^{-4}$ species, with little contribution from $[ss]^{-2}$. The fragments in Fig. 6.2C must therefore result from a two-step process involving non-covalent strand separation ($[ds]^{-4} \rightarrow 2\ [ss]^{-2}$) followed by covalent cleavage of the resulting single strands. This is consistent with the activation time dependence that was reported by Gabelica and De Pauw[11] and interpreted as an effect of entropy: non-covalent dissociation is favored entropically, whereas base loss is a multi-step rearrangement that is entropically disfavored. The fact that the non-covalent pathway was less prevalent for the lower charge state at moderate activation times (Fig. 6.2B vs. D) likely reflects the lower degree of internal Coulombic repulsion in the $[ds]^{-4}$ parent.

Similar to the CAD results obtained for $d(GCGAATTCGC)_2$, CAD spectra for the $[ds]^{-5}$ ions for the duplexes $d(GCGCATGCGC)_2$ and $d(GCAAATTTGC)_2$ exhibited strand separation in a time- and activation voltage-independent fashion (not shown). Results for the $[ds]^{-4}$ charge states demonstrated some sequence dependence, however. (Note that $[ds]^{-4}$ species were so identified in each case because of the spacing of the associated sodium adduct peaks.) Type II fragments were insignificant for all three duplexes when an activation time of 3 ms was employed, and Type I products of single strand cleavage (e.g. $[ss\text{-}GH]^{-2}$) were predominant. As the activation time was increased to 30 ms, Type II products became dominant for $d(GCGCATGCGC)_2$, and products of single-strand fragmentation were far less intense than for $d(GCGAATTCGC)_2$ (in Fig. 6.2D). CAD of $d(GCAAATTTGC)_2$ at 30 ms produced Type II fragments in low proportion, and products deriving from cleavage of single-stranded species were prevalent. These results indicate that at a fixed, moderate activation time and voltage, the

relative intensity of Type II fragments decreased as the GC content of the sequence decreased, in agreement with previously reported trends.[11]

6.3.2 Dissociation of Drug/Duplex Complexes

Small molecules that interact with DNA typically bind Watson-Crick duplexes either at the minor groove or between base pairs (intercalation).[33,34] Minor groove binding generally involves hydrogen bonding interactions, so most compounds that bind in this manner possess several donor/acceptor groups in a curved arrangement to complement the shape of the DNA helix. Intercalation is primarily hydrophobic in nature, and the best intercalators generally consist of fused aromatic rings capable of efficient π-overlap with nucleobase moieties. Because hydrogen bonds tend to survive in the gas phase better than π-π interactions, a recurrent question in ESI-MS studies of DNA duplex behavior is whether these two binding modes can be distinguished with mass spectrometric data.

6.3.2.1 Drug/Duplex Complexes Containing Intercalators

Nogalamycin and daunomycin (structures given in Figure 6.1) are anthracycline antibiotics known to interact with duplex DNA by intercalation.[33] Gas-phase complexes containing these drugs have been shown to mimic binding patterns expected in solution for intercalation between DNA base pairs, with the number of bound ligands being proportional to drug:DNA ratios.[15,35] The electrospray mass spectrum for a solution containing equimolar concentrations of the 10 base pair duplex $d(GCGAATTCGC)_2$ and nogalamycin (Nog) is shown in Figure 6.3A. Intact drug:DNA complexes were observed at 1:1 and 2:1 nogalamycin:duplex stoichiometries in both the -4 and -5 charge states.

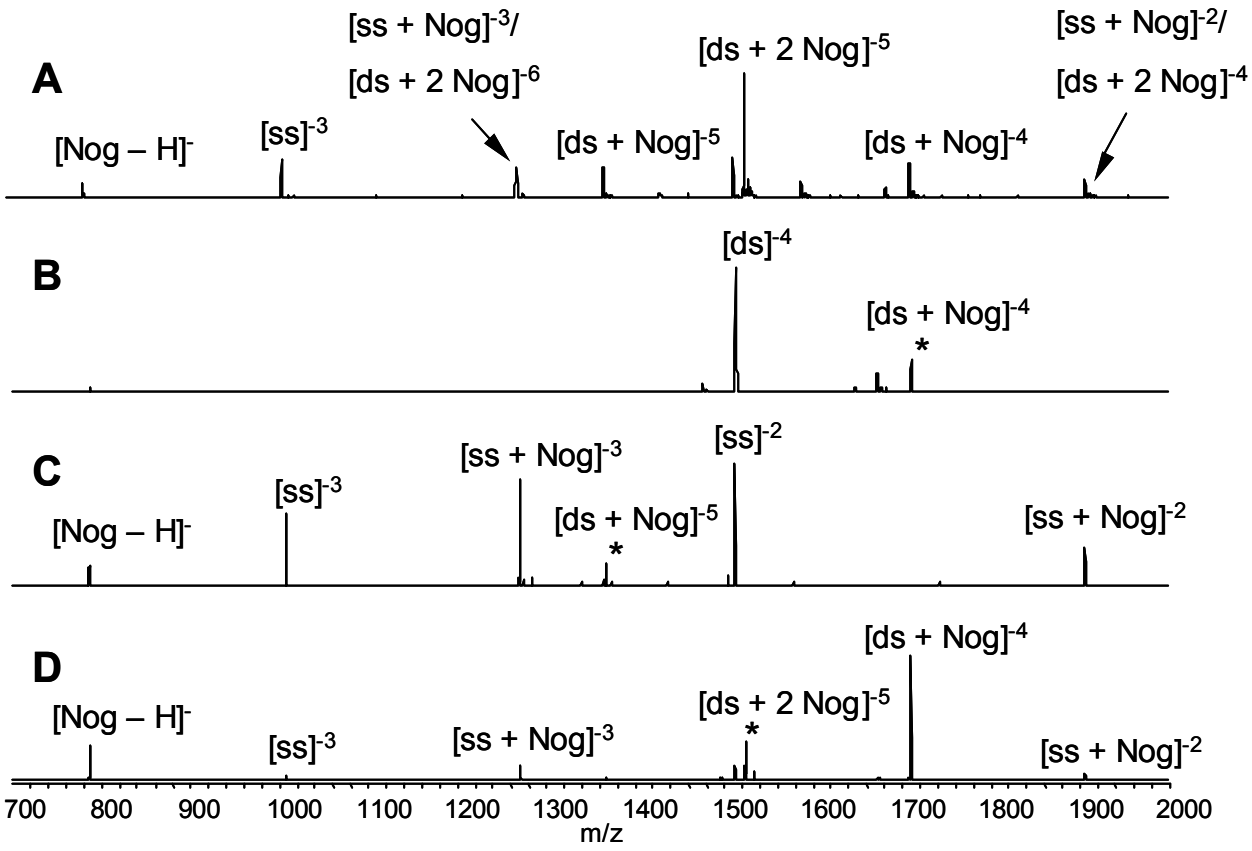


Figure 6.3: ESI and CAD product ion spectra for complexes containing d(GCGAATTCGC)$_2$ and nogalamycin: (A) full scan mass spectrum, (B) CAD of $[ds + Nog]^{-4}$, 0.55V/ 30 ms, (C) CAD of $[ds + Nog]^{-5}$, 0.55V/30 ms, and (D) CAD of $[ds + 2\ Nog]^{-5}$, 0.53V/30 ms. Parent ions are marked *.

(The ion at m/z 1907 is attributed to $[ds + 2\ Nog]^{-4}$ rather than $[ss + Nog]^{-2}$ based on the spacing of the associated sodium adduct peaks.) The CAD spectrum for the 1:1 complex in the -4 charge state (Fig. 6.3B) shows the predominant loss of a neutral drug molecule, with a bare duplex ion remaining at m/z 1513 and no single-stranded species in evidence. In contrast, CAD of the $[1:1]^{-5}$ complex produced intense $[ss]^{-n}$ and $[ss + Nog]^{-n}$ ions (n = 2 or 3) as shown in Fig. 6.3C. (The ion at m/z 1512 was assigned as $[ss]^{-2}$ rather than $[ds]^{-4}$ on the basis of MS^3 experiments, which did not yield the Type II fragments that would have been expected based on MS/MS results in Fig. 6.2D.) These species result from strand separation with retention of drug on one strand. Energy-resolved CAD experiments show that the $[ss]^{-n}$ ions form directly from the precursor and not from secondary loss of drug from $[ss+ Nog]^{-n}$ species, and duplex dissociation is therefore favored over drug ejection for the $[1:1]^{-5}$ complex. Strand separation produces less highly charged species, so the dramatic dependence of fragmentation on the initial charge state is likely a result of Coulombic effects. CAD of the $[2:1]^{-5}$ complex (Fig. 6.3D) gave the $[ds + Nog]^{-4}$ ion as the major product. This process entails charge reduction via ejection of the drug in deprotonated form, which indicates that Coulombic factors affect fragmentation at this stoichiometry as well. The fact that strand separation was not observed suggests that the second nogalamycin molecule binds the duplex more weakly than the first.

To explore the effect of oligonucleotide sequence, CAD spectra were acquired for 1:1 complexes between nogalamycin and $d(GCGCATGCGC)_2$, a duplex with slightly higher GC content. CAD of the $[ds + Nog]^{-4}$ complex with this new sequence (Fig. 6.4A) gave loss of neutral drug, as was observed for the original sequence in Fig. 6.3B. Ejection of the deprotonated drug was the major process for the $[ds + Nog]^{-5}$ parent

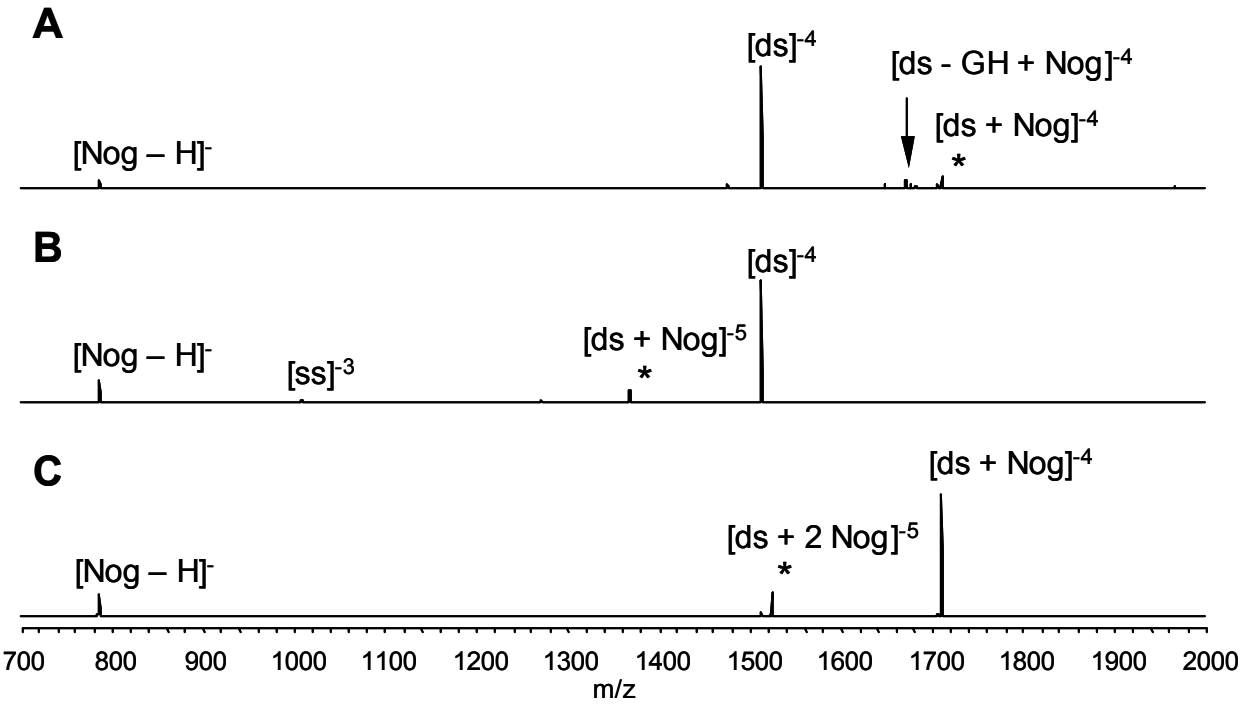


Figure 6.4: CAD product ion spectra for complexes containing $d(GCGCATGCGC)_2$ and nogalamycin: (A) $[ds + Nog]^{-4}$, 0.55V/30 ms, (B) $[ds + Nog]^{-5}$, 0.50V/ 30 ms, and (C) $[ds + 2\ Nog]^{-5}$, 0.50V/30 ms. Parent ions are marked *.

species (Fig. 6.4B), in contrast with the results shown in Fig. 6.3C, in which strand separation was evident. This indicates that a greater number of inter-strand hydrogen bonds can promote the drug ejection pathway over strand separation, as would be expected for duplexes containing greater GC content. Conversely, greater AT content should promote strand separation over drug ejection, and results for complexes with nogalamycin and $d(GCAAATTTGC)_2$ (not shown) provide some support for this. CAD of the $[ds + Nog]^{-5}$ precursor gave a spectrum similar to the one shown in Fig. 6.3C. CAD of the $[ds + 2\ Nog]^{-5}$ precursor (Figure 6.4C), however, produced a greater degree of strand separation than is evident in Fig. 6.3E for the complex with $d(GCGAATTCGC)_2$. While the mode of drug binding to these two duplexes may not be the same, intercalation and minor groove binding cannot be distinguished based on CAD spectra for drug/duplex complexes in high charge states (vide infra).

Experiments conducted with daunomycin indicate that fragmentation pathways can also vary somewhat with the identity of the drug even when the nominal mode of binding (intercalation in this case) is the same. The electrospray mass spectrum for a solution containing $d(GCGAATTCGC)_2$ and daunomycin (Dn) is shown in Figure 6.5A, with the 1:1 and 1:2 duplex:daunomycin complexes labelled. The product ion spectrum for the $[ds + Dn]^{-4}$ complex (Fig. 6.5B) generally resembles the one obtained for the nogalamycin complex (Fig. 6.3B) in that ejection of neutral drug was the predominant process. The $[ds + Dn]^{-5}$ complex, like the $[ds + Nog]^{-5}$ complex, produced significant strand separation (Fig. 6.5C vs. 6.3C). (As in Fig. 6.3C, the ion at m/z 1512 was assigned as $[ss]^{-2}$ rather than $[ds]^{-4}$ on the basis of MS^3 experiments.) Collisional activation of the 1:2 DNA:daunomycin complex in the -5 charge state (Fig. 6.5D) clearly leads to competition between dissociation pathways involving strand separation, neutral drug loss (a route not observed for the 1:2 DNA/nogalamycin complex), and loss of a deprotonated

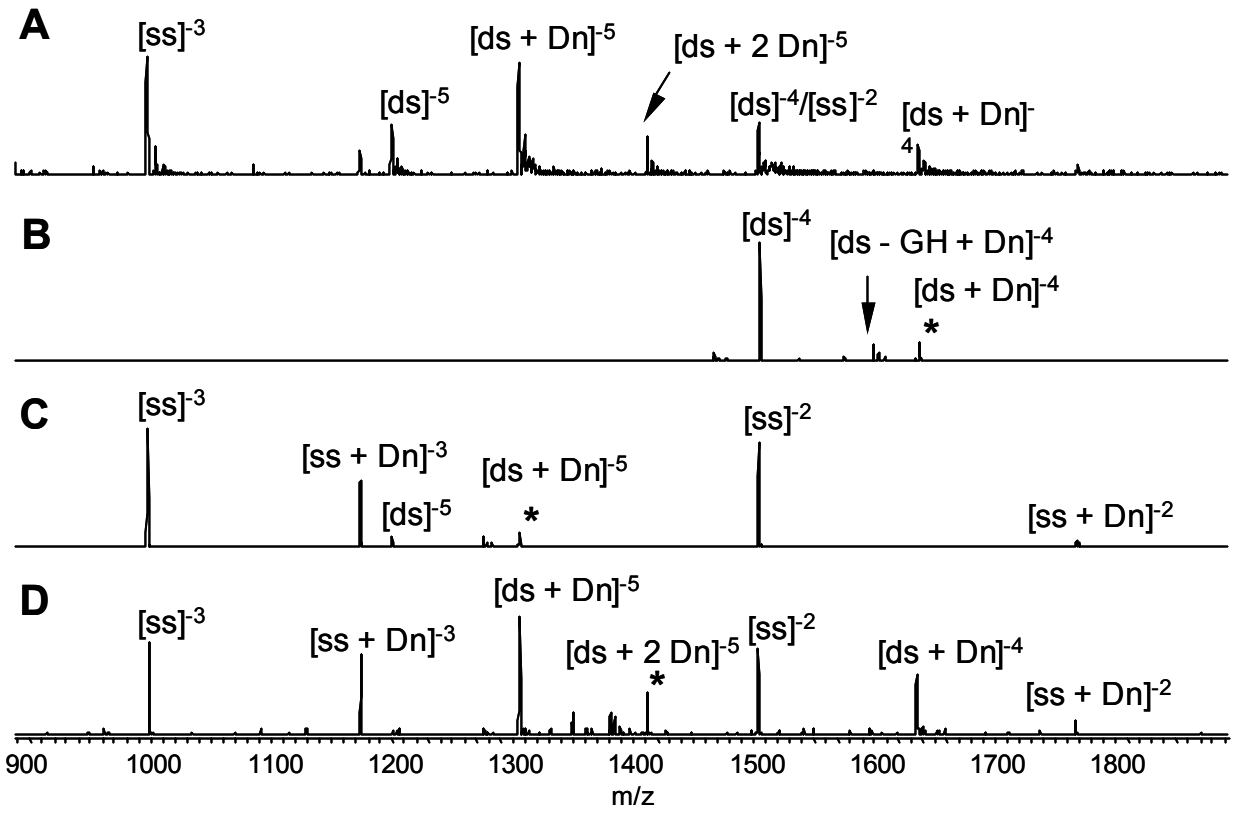


Figure 6.5: ESI and CAD product ion spectra for complexes containing $d(GCGAATTCGC)_2$ and daunomycin: (A) full scan mass spectrum, (B) CAD of $[ds + Dn]^{-4}$, 0.58V/30 ms, (C) CAD of $[ds + Dn]^{-5}$, 0.53V/30 ms, and (D) CAD of $[ds + 2\ Dn]^{-5}$, 0.55V/30 ms. Parent ions are marked *.

daunomycin. The differences apparent in Figs. 6.3D and 6.5D indicate that the observed fragmentation pathways reflect the particular intermolecular interactions present in the complex, and that these interactions can vary for different drugs even when they are known to interact by the same general binding mode. Some common elements emerge nonetheless from the dissociation of duplex/intercalator complexes. The spectra in Figs. 6.3-6.5 give little to no evidence of covalent dissociation, a process observed upon CAD of the duplexes alone, regardless of charge state, sequence, or the identity of the intercalator. In addition, none of these spectra exhibited a strong dependence on activation time and voltage (data not shown). These observations correlate with the analysis presented by Gabelica and DePauw, which postulated time-independent non-covalent fragmentation of gas-phase duplex/intercalator complexes.[11]

6.3.2.2 ***Drug/Duplex Complexes Containing Minor Groove Binders***

To complement data obtained for duplex/intercalator complexes, dissociation experiments were also undertaken with a series of minor groove binding agents. Distamycin (Dist) is an antitumor antibiotic natural product which has been shown to bind the minor groove of B-DNA with an AT base pair specificity.[36] This compound possesses the "classic" crescent-shaped structure common to many minor groove binders, which allows favorable non-covalent interactions between the molecule and DNA bases along the minor groove. Studies investigating distamycin-DNA complexes via electrospray ionization mass spectrometry have been reported showing the generation of gas-phase complexes containing 1:1 and 1:2 DNA:distamycin stoichiometries,[14,16,35] in agreement with solution and crystal structure studies showing the same binding stoichiometries.[37]

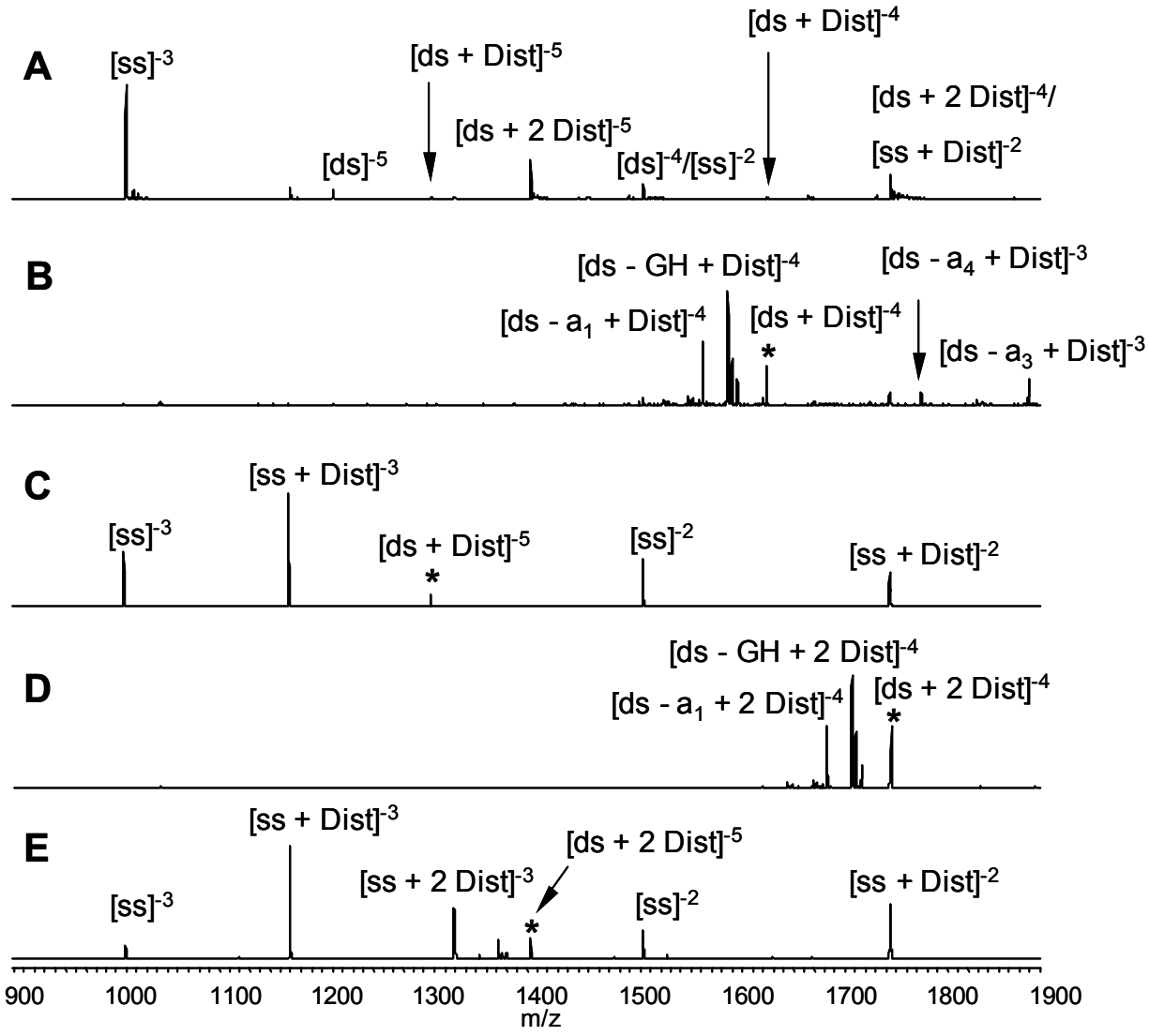


Figure 6.6: ESI and CAD product ion spectra for complexes containing d(GCGAATTCGC)$_2$ and distamycin: (A) full scan mass spectrum, (B) CAD of [ds + Dist]$^{-4}$; 0.63V/30 ms, (C) CAD of [ds + Dist]$^{-5}$, 0.55V/30 ms (D) CAD of [ds + 2 Dist]$^{-4}$, 0.6V/30 ms and (E) CAD of [ds + 2 Dist]$^{-5}$, 0.50V/30 ms. Parent ions are marked *.

Figure 6.6A shows a representative mass spectrum for a solution containing equimolar concentrations of $d(GCGAATTCGC)_2$ and distamycin (Dist). Ions corresponding to complexes of the type $[ds + Dist]^{-n}$ and $[ds + 2\ Dist]^{-n}$ were observed in both the -5 and -4 charge states. (The peak at m/z 1754 is attributed primarily to $[ds + 2\ Dist]^{-4}$ rather than $[ss + Dist]^{-2}$ based on the spacing of the associated sodium adduct peaks.) CAD of the $[ds + Dist]^{-4}$ complex (Fig. 6.6B) produced a variety of Type II fragments (e.g. $[ds - GH]^{-4}$, $[ds - a_1 + Dist]^{-4}$). Little evidence of strand separation or drug loss was observed. The CAD spectrum for the -5 charge state complex (Fig. 6.6B) shows a dramatic difference in dissociation pathway, with strand separation predominant over drug or base loss. Relative to the intercalators, the lack of drug loss for either the –4 or –5 charge states of the distamycin/duplex complexes may be attributed to the hydrogen bonding interactions between the amide nitrogens on the drug and the nucleobases exposed at the floor of the minor groove, and perhaps also to the strong gas-phase electrostatic interactions between the permanently-charged guanidinium group of distamycin and the DNA backbone phosphates. The marked differences in dissociation pathways for these complexes once again illustrate the effect of Coulombic interactions upon complex stability.

CAD spectra for the $[ds + 2\ Dist]^{-4}$ and $[ds + 2\ Dist]^{-5}$ species are shown in Fig. 6.6D and E, respectively. Like $[ds + Dist]^{-4}$, $[ds + 2\ Dist]^{-4}$ gave Type II products of covalent dissociation almost exclusively (compare Figs. 6.6B and D), and like $[ds + Dist]^{-5}$, $[ds + 2\ Dist]^{-5}$ underwent only strand separation (compare Figs. 6.6C and E). These results closely parallel the charge state-dependent trends obtained for the 1:1 complexes. The fact that little evidence of drug loss was observed for the 1:2 complexes indicates that the second distamycin molecule binds virtually as strongly as the first. This contrasts with data for the nogalamycin complexes, where the $[ds + 2\ Nog]^{-5}$ precursor

dissociated primarily by ejection of deprotonated nogalamycin. It is, however, consistent with the formation of a "head-to-tail" dimer in which both distamycin molecules bind via comparable interactions with the duplex; this binding mode has been observed in previous NMR studies.[38]

Distamycin also formed 1:1 and 1:2 complexes with $d(GCGCATGCGC)_2$ and $d(GCAAATTTGC)_2$, and as in Fig. 6.6A the 1:2 stoichiometry was favored based on its greater signal intensity. The CAD spectrum of the $[ds + Dist]^{-4}$ complex with $d(GCAAATTTGC)_2$, acquired with an activation time of 30 ms, is shown in Fig. 6.7B. This precursor gave $[ss]^{-2}$ and $[ss + Dist]^{-2}$ fragments via strand separation in addition to several Type II fragments via covalent cleavage, again showing that the base composition of the DNA duplex influences the preferred dissociation pathways. (In this case m/z 1512 could not be distinguished as either $[ss]^{-2}$ or $[ds]^{-4}$ because as described above, CAD of $[d(GCAAATTTGC)_2]^{-4}$ gave only products deriving from cleavage of single-strands and no characteristic Type II fragments.) Unlike the lower charge states of the duplex/intercalator species, however, $[ds + drug]^{-4}$ parent ions for duplex/minor groove binder complexes gave fragmentation patterns that were strongly dependent on the activation conditions. This is illustrated for the complex with $d(GCAAATTTGC)_2$ by comparing Fig. 6.7B with Fig. 6.7A and C: the relative intensity of the $[ss]^{-2}$ and $[ss + Dist]^{-2}$ ions decrease as activation time increases and activation voltage decreases. This is consistent with the CAD results for the drug-free duplex species shown in Fig. 6.2C-E, and it also bears out the prediction that dissociation data for DNA/minor groove binder complexes should be dependent on activation regime (and therefore activation time).[11]

Figure 6.8 shows CAD spectra for $[ds + drug]^{-4}$ parent species produced by ESI of $d(GCGAATTCGC)_2$ and the minor groove binders netropsin, DAPI, and Hoechst 33342. The spectra for the netropsin and the DAPI complexes (Fig. 6.8A and B) are dominated

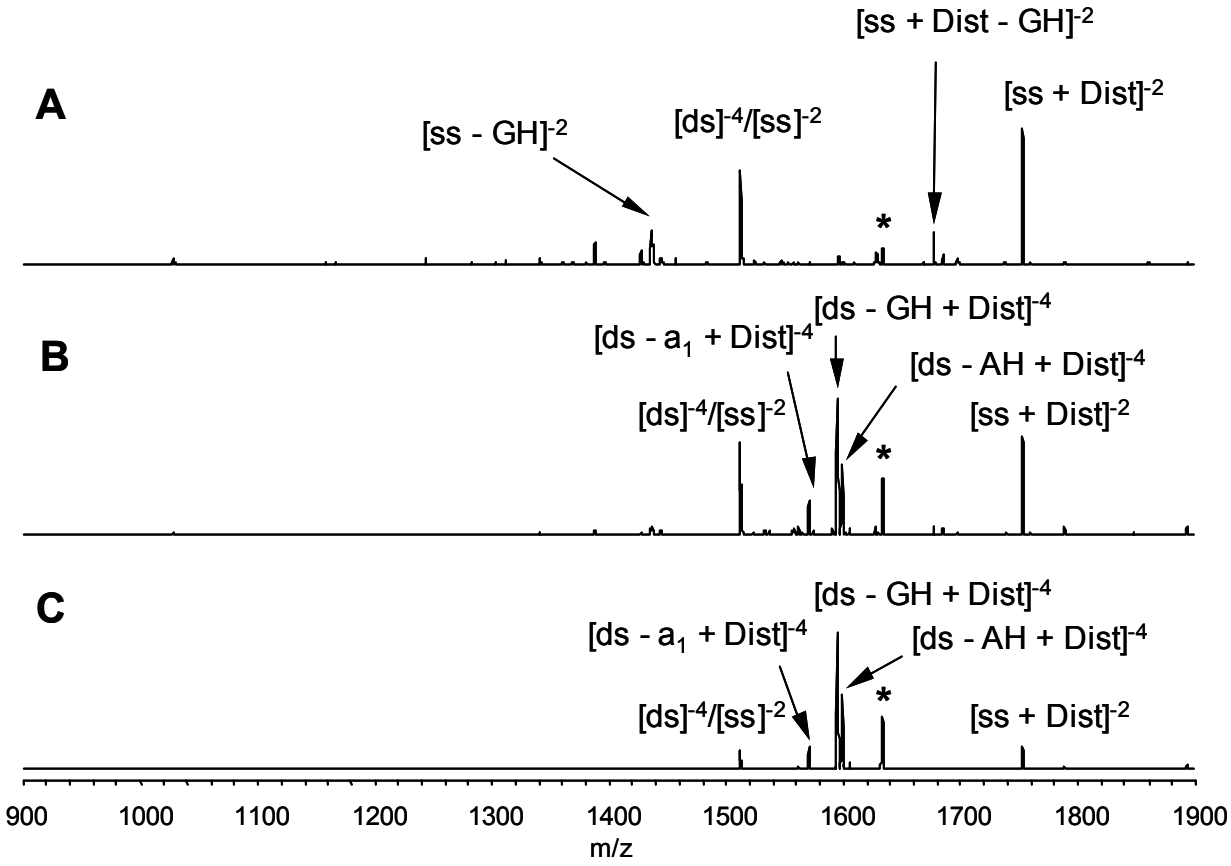


Figure 6.7: CAD product ion spectra for $[d(GCAAATTTGC)_2 + Dist]^{-4}$. (A) 1.2V/3 ms, (B) 0.60V/30 ms and (C) 0.48V/300 ms. The parent ion is marked *.

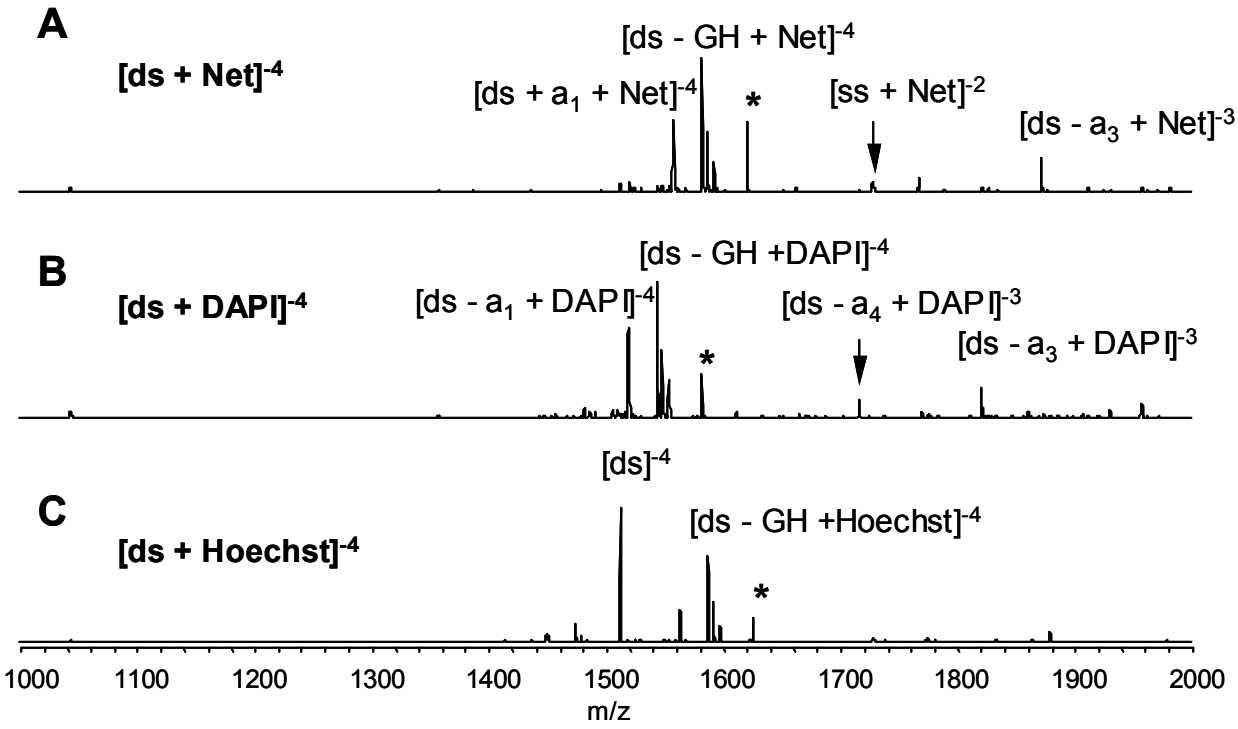


Figure 6.8: CAD product ion spectra for $[d(GCGAATTCGC)_2 + drug]^{-4}$ complexes with (A) netropsin, 0.65V/30 ms, (B) DAPI, 0.63V/30 ms and (C) Hoechst 33342 0.63V/30 ms. Parent ions are marked *.

by products of covalent cleavage, like the spectrum obtained for the distamycin complex of this sequence (Fig. 6.6B). For the complex with Hoechst 33342, however, ejection of the neutral drug was a competitive process, as indicated by the presence of the $[ds]^{-4}$ ion in Fig. 6.8C. This result could indicate that Hoechst 33342 binds this duplex less avidly than the other drugs in the gas phase; its structure contains fewer hydrogen bond donors than distamycin and netropsin, and fewer charged groups than netropsin and DAPI. However, solution studies have prompted speculation that Hoechst 33258, a close analog of Hoechst 33342, may bind duplex DNA both in the minor groove and to a lesser extent by non-classical intercalation.[39-42] The results presented above for nogalamycin and daunomycin shows that the drug ejection pathway is favored for duplex/intercalator complexes in low charge states, so the spectrum in Fig. 6.8C could also be interpreted as evidence of multiple binding modes for Hoechst 33342. The proportion of drug ejection vs. covalent dissociation decreased as activation time increased and activation voltage decreased (data not shown), loss of drug remained a major process at long activation times.

Norfloxacin belongs to the fluoroquinolone family of antibiotics, which exhibit potent antibacterial activity through the inhibition of DNA gyrase.[43-48] The primary binding target for these compounds is thought to be the DNA substrate for gyrase rather than the protein itself, but the specific DNA binding mode for the fluoroquinolones remains controversial. Recent spectroscopic and molecular modeling investigations suggest that norfloxacin may interact with duplex DNA by multiple binding modes, including both groove and non-classic intercalative interactions, and some evidence for a GC base pair preference has been reported.[49-52] To our knowledge the gas-phase interactions between oligonucleotides and norfloxacin have not been studied. ESI-MS and CAD were therefore used to evaluate DNA-norfloxacin interactions, and the results

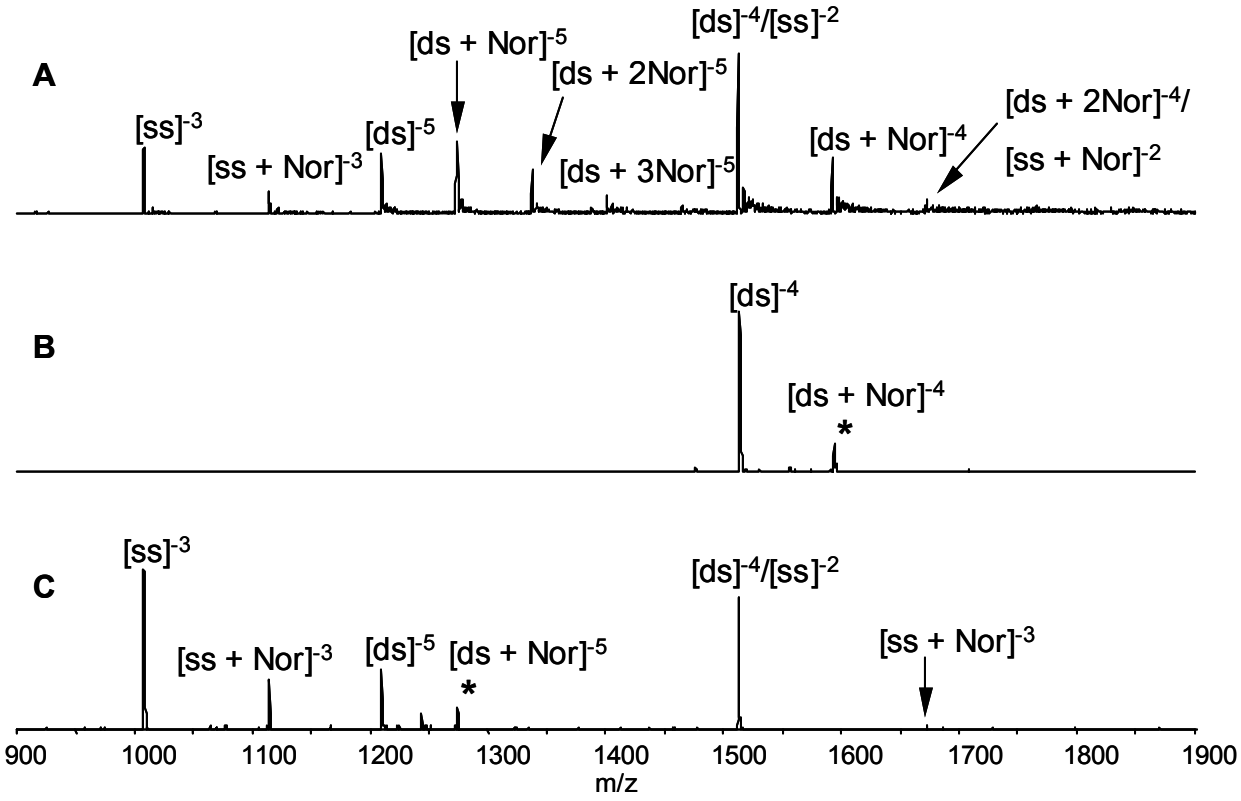


Figure 6.9: ESI and CAD product ion spectra for complexes containing $d(GCGAATTCGC)_2$ and norfloxacin: (A) full scan mass spectrum, (B) CAD of $[ds + Nor]^{-4}$, 0.53V/30 ms and (C) CAD of $[ds + Nor]^{-5}$, 0.53V/30 ms. Parent ions are marked *.

were interpreted in light of results obtained for the minor groove binding and intercalating drugs discussed above.

The electrospray mass spectrum for solutions containing $d(GCGAATTCGC)_2$ and norfloxacin is shown in Figure 6.9A. Ions corresponding to -4 and -5 charge state complexes containing from 1:1 to 3:1 norfloxacin:DNA ratios are detected, with a binding pattern and distribution reminiscent of gas-phase DNA-drug complexes containing intercalators as in Figs. 6.3A and 6.5A. The CAD product ion spectra for the -4 and -5 charge state 1:1 complexes (Fig. 6.9B and C) produce results similar to those observed for nogalamycin and daunomycin, namely preferential loss of one neutral drug from the -4 charge state parent ion, and strand dissociation/drug loss from the –5 charge state precursor. These results do not strictly preclude interactions with the minor groove, but in comparison with results for minor groove binders, particularly distamycin, netropsin, and DAPI, they do seem to rule out a binding modes that might involve an extensive network of hydrogen bonds.

6.4 CONCLUSIONS

The results presented above demonstrate the dramatic effect of charge state upon the dissociation of gas-phase DNA-DNA and drug-DNA complexes. Generally, duplex strand separation predominates for higher charge states. This characteristic "unzipping" is likely due to complex destabilization caused by internal Coulombic repulsion, and was found to be relatively insensitive to activation time and voltage. The DNA sequence and/or the structure of the drug can counteract this tendency to some extent, indicating that the phenomenon reflects the relative strength of both the DNA-DNA and drug-DNA interactions. For lower charge states, ejection of either a neutral or deprotonated drug molecule predominated for DNA/intercalator complexes, and this was shown to be nearly

independent of activation time. For complexes with minor groove binders, covalent cleavage was generally favored over drug ejection, although the competition between these processes was strongly dependent on activation time. The dissociation pathways for the lower charge state complexes are probably more reflective of specific drug-DNA interactions because Coulombic effects are less marked for these precursors. Thus, CAD analysis of the lower charge state complexes is more appropriate in evaluating gas-phase drug-DNA binding in the absence of the competing duplex strand separation pathway.

6.5 REFERENCES

(1) Smith, D. L.; Deng, Y.; Zhang, Z. *J. Mass Spectrom.* **1997**, *32*, 135-146.

(2) Smith, R. D.; Bruce, J. E.; Wu, Q.; Lei, Q. P. *Chem. Soc. Rev.* **1997**, *26*, 191-202.

(3) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(4) Loo, J. A.; Loo, R. R. O. In *Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications*; Cole, R. B., Ed.; John Wiley & Sons: New York, 1997; pp 385-419.

(5) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(6) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, *120*, 9605-9613.

(7) Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. *Rapid Commun. in Mass Spectrom.* **2000**, *14*, 464-467.

(8) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(9) Gabelica, V.; De Pauw, E. *J. Mass Spectrom.* **2001**, *36*, 397-402.

(10) Gabelica, V.; De Pauw, E. *Int. J. Mass Spectrom.* **2002**, *219*, 151-159.

(11) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

(12) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.

(13) Goodlett, D. R.; Camp, D. G., II; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, *22*, 181-183.

(14) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(15) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(16) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(17) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.

(18) Carrasco, C.; Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E.; Garbay-Jaureguiberry, C.; Roques, B.; Wilson, W. D.; Chaires, J. B.; Waring, M. J.; Bailly, C. *ChemBioChem* **2002**, *3*, 1235-1241.

(19) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(20) Guittat, L.; Alberti, P.; Rosu, F.; Van Miert, S.; Thetiot, E.; Pieters, L.; Gabelica, V.; De Pauw, E.; Ottaviani, A.; Riou, J.-F.; Mergny, J.-L. *Biochimie* **2003**, *85*, 535-547.

(21) Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

(22) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 10129-10133.

(23) Hofstadler, S. A.; Sannes-Lowery, K. A.; Crooke, S. T.; Ecker, D. J.; Sasmor, H.; Manalili, S.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 3436-3440.

(24) Griffey, R. H.; Sannes-Lowery, K. A.; Drader, J. J.; Mohan, V.; Swayze, E. E.; Hofstadler, S. A. *J. Am. Chem. Soc.* **2000**, *122*, 9933-9938.

(25) Sannes-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264-271.

(26) Sannes-Lowery, K. A.; Drader, J. J.; Griffey, R. H.; Hofstadler, S. A. *Proc. SPIE-Int. Soc. Opt. Eng.* **2001**, *4264*, 27-36.

(27) Hogan, J. M.; McLuckey, S. A. *J. Mass Spectrom.* **2003**, *38*, 245-256.

(28) Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal. Chem.* **2001**, *73*, 3274-3281.

(29) Newton, K. A.; Chrisman, P. A.; Reid, G. E.; Wells, J. M.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2001**, *212*, 359-376.

(30) Wells, J. M.; Stephenson, J. L.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2000**, *203*, A1-A9.

(31) McLuckey, S. A.; Vaidyanathan, G. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *162*, 1-16.

(32) Keller, K. M.; Brodbelt, J. S. *Anal. Biochem.* **2004**, *326*, 200-210.

(33) Kallenbach, N. R., Ed. *Chemistry and Physics of DNA-Ligand Interactions*; Adenine Press: Schenectady, NY, 1990.

(34) Propst, C. L.; Perun, T. J., Eds. *Nucleic Acid Targeted Drug Design*; M. Dekker: New York, 1992.

(35) Gupta, R.; Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 2472-2480.

(36) Zimmer, C.; Waehnert, U. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31-112.

(37) Uytterhoeven, K.; Sponer, J.; Van Meervelt, L. *Eur. J. Biochem.* **2002**, *269*, 2868-2877.

(38) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 5723-5727.

(39) Bailly, C.; Henichart, J. P.; Colson, P.; Houssier, C. *J. Mol. Recognit.* **1992**, *5*, 155-171.

(40) Bailly, C.; Colson, P.; Henichart, J. P.; Houssier, C. *Nucleic Acids Res.* **1993**, *21*, 3705-3709.

(41) Colson, P.; Houssier, C.; Bailly, C. *J. Biomol. Struct. Dyn.* **1995**, *13*, 351-366.

(42) Steinmetzer, K.; Reinert, K. E. *J. Biomol. Struct. Dyn.* **1998**, *15*, 779-791.

(43) Shen, L. L.; Baranowski, J.; Pernet, A. G. *Biochemistry* **1989**, *28*, 3879-3885.

(44) Shen, L. L.; Mitscher, L. A.; Sharma, P. N.; O'Donnell, T. J.; Chu, D. W.; Cooper, C. S.; Rosen, T.; Pernet, A. G. *Biochemistry* **1989**, *28*, 3886-3894.

(45) Marians, K. J.; Hiasa, H. *J. Biol. Chem.* **1997**, *272*, 9401-9409.

(46) Khodursky, A. B.; Cozzarelli, N. R. *J. Biol. Chem.* **1998**, *273*, 27668-27677.

(47) Kampranis, S. C.; Maxwell, A. *J. Biol. Chem.* **1998**, *273*.

(48) Kwok, Y.; Zeng, Q.; Hurley, L. H. *J. Biol. Chem.***1999**, *274*, 17226-17235.

(49) Son, G. S.; Yeo, J.-A.; Kim, M.-S.; Kim, S. K.; Holmen, A.; Akerman, B.; Norden, B. *J. Am. Chem. Soc.* **1998**, *120*, 6451-6457.

(50) Son, G. S.; Yeo, J.-A.; Kim, J.-M.; Kim, S. K.; Moon, H. R.; Nam, W. *Biophys. Chem.* **1998**, *74*, 225-236.

(51) Lee, E.; Yeo, J.-A.; Cho, C. B.; Lee, G. J.; Han, S. W.; Kim, S. K. *Eur. J. Biochem.* **2000**, *267*, 6018-6024.

(52) Lee, H. M.; Kim, J.-K.; Kim, S. K. *J. Biomol. Struct. Dyn.* **2002**, *19*, 1083-1091.

Chapter 7: Electrospray Ionization of Nucleic Acid Aptamer/Small Molecule Complexes for Screening Aptamer Selectivity

7.1 INTRODUCTION

Aptamers are nucleic acids (strands of DNA or RNA) that are capable of recognizing target molecules with high specificity and affinity.[1-5] Aptamers are generally "unnatural" sequences, isolated not from biological sources but rather by *in vitro* selection from combinatorial libraries.[1-7] The selection process typically begins with a large pool of DNA or RNA (10^{13}-10^{15} randomized sequences), which is gradually reduced in complexity by multiple cycles of affinity chromatography and PCR amplification. The sequences that survive this process can bind their target analytes with affinities (K_d) ranging from 10^{-5} to 10^{-13} M and discriminate between molecules that differ by as little as a single methyl group.[3,4] Because of these recognition properties, aptamers can be employed as sensitive diagnostic agents,[5] biomedical research tools,[3,8] and even as therapeutics.[9] On a more fundamental level, aptamers readily lend themselves to structural analysis by NMR spectroscopy because of their modest size. Several three-dimensional binding motifs later found in biologically relevant systems were originally identified during structural characterization of aptamer/ligand interactions.[4]

Electrospray ionization mass spectrometry (ESI-MS) has been used to assess binding stoichiometry and relative binding affinity in numerous molecular recognition processes, from cation/crown ether interactions[10] to drug/receptor interactions.[11,12,13] The electrospray process is gentle and can preserve, at least to some extent, the delicate

biomolecular architecture involved in many protein/nucleic acid,[11,12,14] nucleic acid/nucleic acid, [12] and nucleic acid/ligand interactions,[12,13] which suggests that aptamer/ligand interactions might survive as well. If so, mass spectrometry could provide a tool for rapidly screening aptamer selectivity, and might possibly accelerate the process of identifying high-affinity sequences during selection. Recently, a gas-phase complex between the transcription factor NF-kappaB and its RNA aptamer has been reported,[15,16] illustrating that aptamer interactions with large, pre-organized target molecules can be detected by ESI-MS. To our knowledge, however, no such reports involving small molecule ligands have appeared. We therefore selected several previously characterized aptamers that recognize small molecules as case studies for ESI-MS analysis, including aptamers for tobramycin,[17,18] adenosine triphosphate (ATP),[19,20] flavin mononucleotide (FMN),[21-23] and theophylline.[24-26] These receptors bind their targets with affinities ranging from 6 μM to 12 nM through specific intermolecular interactions that have been identified in independent structural studies. ESI experiments were conducted for each aptamer in the presence of several ligands, and in some cases parallel solution assays were conducted to provide data for comparison. Our results indicate that only limited correlations can be drawn between the gas- and solution-phase behavior of aptamer/ligand complexes.

7.2 EXPERIMENTAL

7.2.1 Materials

The 27-residue DNA aptamer that binds adenosine, AMP, and ATP (5'-ACCTGGGGGAGTATTGCGGAGGAAGGT-3')[19] was purchased from Invitrogen (Carlsbad, CA). RNA aptamers that bind tobramycin (5'-GGGACUUGGUUUA-

GGUAAUGAGUCCCU-3')[18], FMN (5'-GGCGUGUAGGAUAUGCUUCGGCAGAA-GGACACGCC-3')[22] and theophylline (5'-GGCGAUACCAGCCGAAAGGCCCUUG-GCAGCGUC-3') were chemically synthesized on an Expedite 8909 synthesizer (Perkin-Elmer, Norwalk, CT) using standard RNA phosphoramidite reagents purchased from Glen Research (Sterling, VA). Adenosine triphosphate (ATP) was purchased from Epicentre (Madison, WI), and all other ligands (structures given in Figure 7.1) were purchased from Sigma (St. Louis, MO). Gentamicin (Sigma) was obtained as a mixture of the C1, C2, and C2A variants and was used as received. $MnCl_2$ was purchased from Aldrich (Milwaukee, WI).

7.2.2 Mass Spectrometry

All mass spectrometry experiments were performed on an LCQ Duo quadrupole ion trap mass spectrometer (ThermoQuest, San Jose, CA) using a heated capillary temperature of 80-90°C to minimize decomposition in the electrospray interface. Aptamer solutions contained 20 μM oligonucleotide and 50 mM ammonium acetate in 10% isopropanol/water (ATP aptamer), or 2 μM oligonucleotide and 50 mM ammonium acetate in 30% isopropanol/water (tobramycin, FMN, and theophylline aptamers). Ligands (1-5 eq.) and/or divalent metal salts (2-5 eq.) were also added as appropriate. The final pH of these solutions was approximately 6.5. Electrospray experiments were conducted in negative ion mode (-3.4 kV) by direct infusion (3 μl/min). For experiments with tobramycin and other aminoglycosides, the usual fused silica infusion line was replaced with deactivated fused silica to prevent sample carryover. Collision-activated dissociation (CAD) experiments were performed on selected precursors using activation voltages sufficient to reduce the parent ion intensity by 80-90%. One to two hundred scans were averaged to produce each mass spectrum.

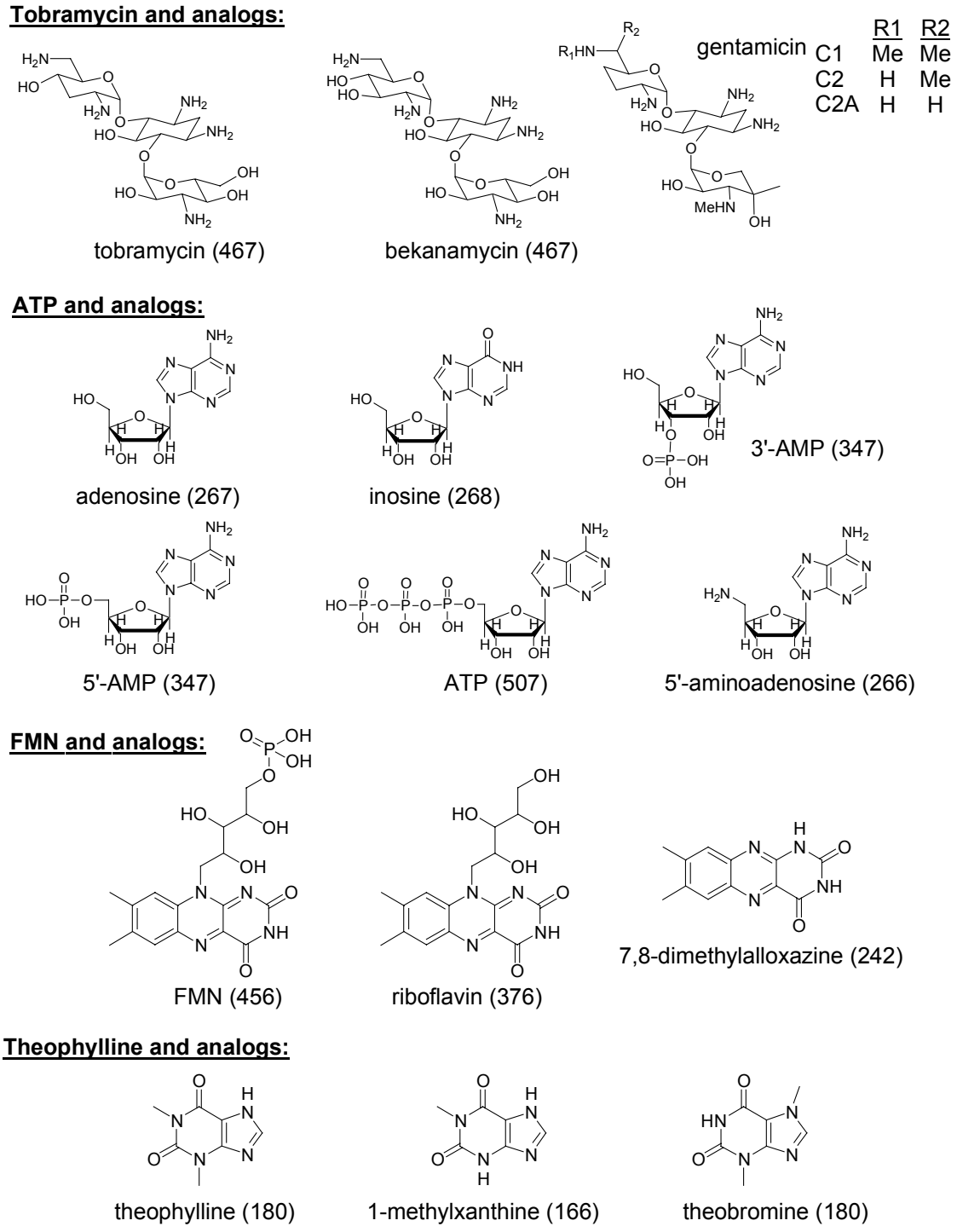


Figure 7.1: Structures of ligands used in ESI-MS experiments. Molecular weights (Da) are given in parentheses; common abbreviations are also indicated.

7.2.3 Column Binding Assays

Buffers used in selection experiments typically employ both monovalent and divalent metal salts (usually NaCl and $MgCl_2$) at approximate concentrations of 100-300 mM and 1-5 mM, respectively. Such high concentrations of involatile salt strongly suppress ionization in electrospray experiments, however, and can also clog the sampling orifice of the mass spectrometer. For this reason, the primary buffer salt used in the electrospray solutions was ammonium acetate, which is relatively volatile and therefore better tolerated in ESI-MS experiments. Furthermore, when ESI mass spectra indicated that divalent metal salts were required to mediate aptamer/ligand binding (as indicated by the presence of aptamer/ligand complexes incorporating adventitious cations), $MnCl_2$ was employed to provide aptamer/metal/ligand complexes with unambiguous m/z values.

Aptamer/ligand binding is often quite sensitive to buffer conditions, however, and relatively minor changes in the composition and/or pH of the original selection buffer can reduce binding efficiency or possibly change relative binding affinities. Ligand competition experiments were therefore conducted for both the ATP and FMN aptamers using a buffer containing 50 mM ammonium acetate and 5 mm $MnCl_2$, which approximates the solution conditions used in the electrospray experiments.

7.2.3.1 ATP Binding Assays

Competition assays were conducted using the column binding procedure originally reported by Huizenga and Szostak.[19] Briefly, the ATP aptamer was radiolabeled at the 5' end using T4 polynucleotide kinase purchased from New England Biolabs (Beverly MA) and [γ -^{32}P] ATP (7000 ci/mmol) purchased from MP Biomedicals (Aurora, OH). After gel purification, 40-60 pmol of γ-labeled ATP aptamer was denatured at 72°C (2 min.), then cooled to room temperature (20-30 min.) and incubated

at room temperature (30 min.) with ATP agarose purchased from Sigma (St. Louis, MO). The pre-equilibrated agarose was then loaded onto a chromatography column purchased from Bio-Rad Laboratories (Hercules, CA) and washed with three column volumes of binding buffer. In these studies the standard buffer (300 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris, pH 7.6)[19] was replaced with a buffer comprised of 50 mM ammonium acetate, 5 mM $MnCl_2$ at pH 6.5. The DNA aptamer was then eluted with 1-3 mM ligand (adenosine, inosine, 3'-AMP, 5'-AMP, ATP, or 5'-aminoadenosine) in binding buffer. The amount of DNA retained on the column and in the wash and elute volumes was determined by scintillation counting, and the percentage of eluted DNA aptamer was determined.

7.2.3.2 FMN Binding Assays

Competition assays were conducted as previously described.[21] FMN-agarose was prepared by conjugating FMN with cyanogen bromide-activated agarose (Sigma, St. Louis MO) at 5 mg/mL. The RNA aptamer was radiolabeled at the 5' end using T4 polynucleotide kinase purchased from New England Biolabs (Beverly, MA) and [γ -^{32}P] ATP 7000ci/mmol purchased from MP Biomedicals (Aurora, OH). Following gel purification, 40-60pmol of γ-labeled FMN aptamer was denatured at 72°C (2 min.), then cooled to room temperature (20-30 min.) and incubated at room temperature with FMN agarose (30 min.). The pre-equilibrated agarose was then loaded onto a chromatography column purchased from Bio-Rad Laboratories (Hercules CA) and washed with three column volumes of binding buffer. In these studies the standard buffer (250 mM NaCl, 5 mM $MgCl_2$, 50 mM Tris, pH 7.6)[21] was replaced with a buffer comprised of 50 mM ammonium acetate, 5 mM $MnCl_2$ at pH 6.5. The RNA aptamer was then eluted with 1-3 mM ligand (riboflavin, FMN, and 7,8-dimethylalloxazine) in binding buffer. The

amount of RNA retained on the column and in the wash and elute volumes was determined by scintillation counting, and the percentage of eluted RNA aptamer was determined.

7.3 RESULTS AND DISCUSSION

To evaluate mass spectrometry as a tool for screening molecular recognition and discrimination by DNA and RNA aptamers, ESI-MS experiments and were conducted for each of four previously characterized aptamers in the presence of several ligands. The relative binding affinities determined by ESI-MS were compared to trends established by traditional solution methods. For the ATP and FMN aptamers, solution binding trends were determined experimentally, and for the tobramycin and theophylline aptamers trends reported in previous literature[17,18,24] were used for comparison. In some cases collision-activated dissociation (CAD) was utilized to further assess the relative stability of the observed gas-phase complexes.

7.3.1 Tobramycin Aptamer/Ligand Interactions

At pH 7.4, the 27-residue RNA aptamer for tobramycin[18] binds one ligand in a high-affinity site (K_d = 12 nM) and can also bind a second tobramycin molecule in a lower affinity site (13 μM).[17] The related aminoglycoside bekanamycin binds with similar affinity, but gentamicin, which has a different pattern of substituents on its two sugar rings, binds much less efficiently.[18] The 1:1 complex with tobramycin has been studied by NMR at pH 6.8,[18] and the resulting three-dimensional structure (PDB code 2TOB)[27] reveals a distorted hairpin comprised of a 6 base-pair stem and a 14-residue loop region, which provides a deep pocket for ligand binding. The size of the pocket and a network of hydrogen bonds between protonated amino groups on the ligand and

backbone phosphates on the aptamer provide the basis for high affinity binding by tobramycin and bekanamycin, but preclude effective interactions with gentamicin.

Full scan ESI-MS spectra acquired for solutions containing this aptamer and various aminoglycosides are shown in Figure 7.2. Under the conditions used in these experiments (50 mM ammonium acetate, 30% isopropanol/water) the aptamer produced ions at m/z 1443 and 1732, corresponding to the –6 and –5 charge states (Fig. 7.2A). In the presence of one equivalent of either tobramycin or bekanamycin, the dominant species were 1:1 aptamer:ligand complexes, with 1:2 complexes also visible at much lower intensity (Fig. 7.2B and C). Signals for free RNA were barely detectable in these spectra. In the presence of gentamicin, however, peaks for free RNA were dominant and peaks for 1:1 complexes were relatively insignificant (Fig. 7.2D). These results correlate exactly with the known trend in solution binding[18] and therefore indicate that the gas-phase data reflect specific binding behavior. Titration of the aptamer with increasing amounts of tobramycin demonstrated that 1:1 and 1:2 aptamer:ligand complexes form sequentially (Fig. 7.3), which is consistent with the wide difference in their solution binding constants and provides further evidence of the specific nature of the interactions reflected in the MS data.

CAD experiments were performed to assess the kinetic stability of these complexes, and representative results are shown in Figure 7.4. The –5 charge state of the 1:1 complex with tobramycin dissociated by covalent cleavage (loss of nucleobase moieties or H_2O) rather than non-covalent loss of ligand (Fig. 7.4A). Competition between covalent and non-covalent cleavage processes has been observed in previous CAD studies of both DNA duplexes and duplex/drug complexes in the quadrupole ion trap,[28-31] and both the activation conditions[30] and initial charge state of the precursor[31] have been shown to affect this competition. In experiments designed to promote non-

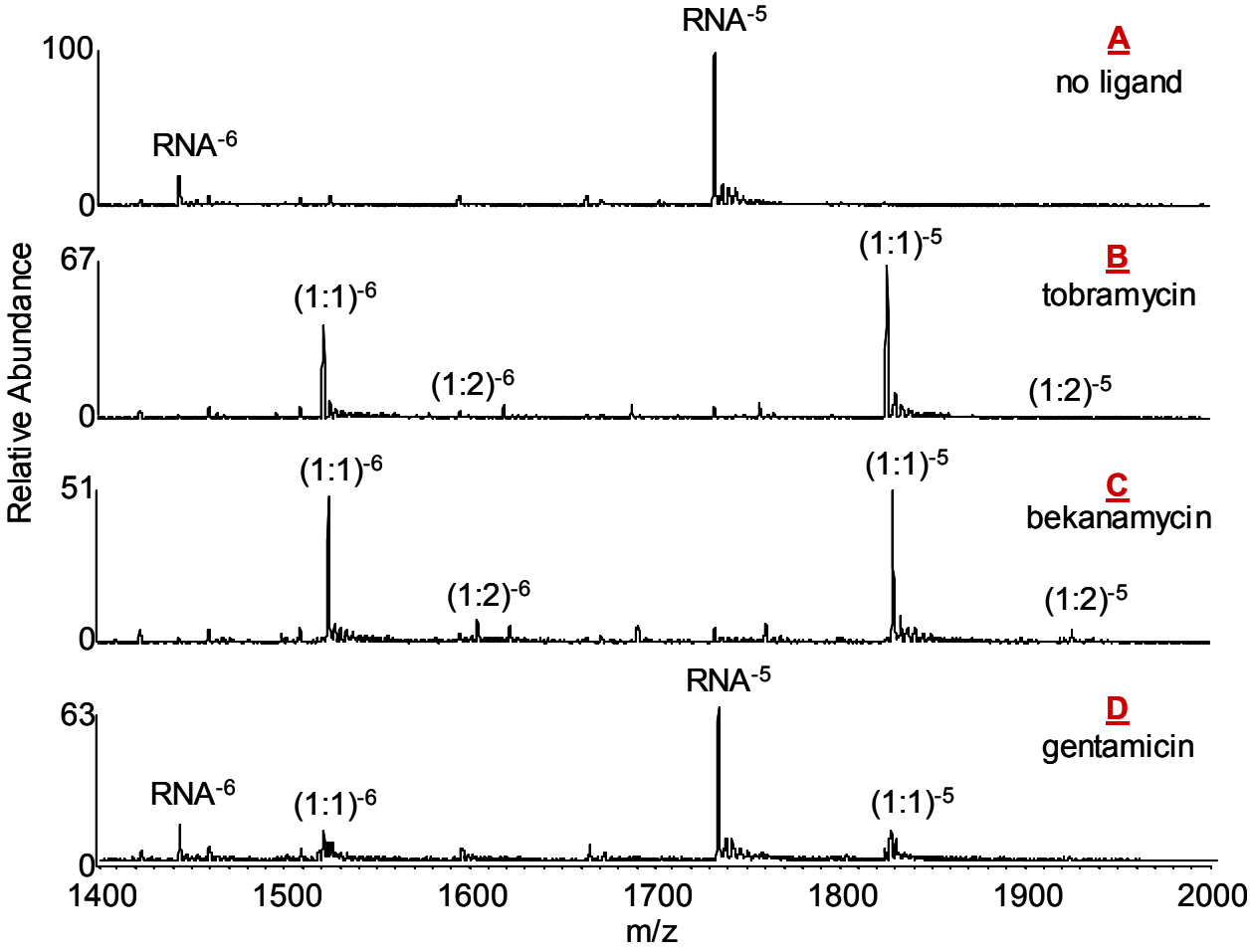


Figure 7.2: ESI spectra for samples containing the tobramycin aptamer (RNA). A) Tobramycin aptamer, B) aptamer + 1 eq. tobramycin, C) aptamer + 1 eq. bekanamycin, and D) aptamer + 1 eq. gentamicin (mixture of isomers). All solutions contained 2 μM RNA and 50 mM NH_4OAc in 30% IPA/H_2O.

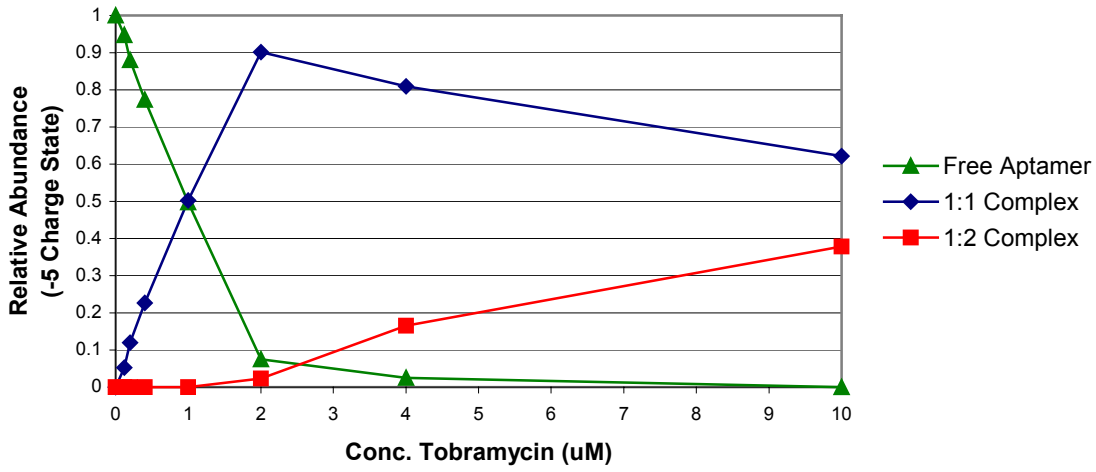


Figure 7.3: Titration of tobramycin aptamer (2 μM) with tobramycin. Plot of tobramycin concentration vs. relative abundance of ions in the -5 charge state.

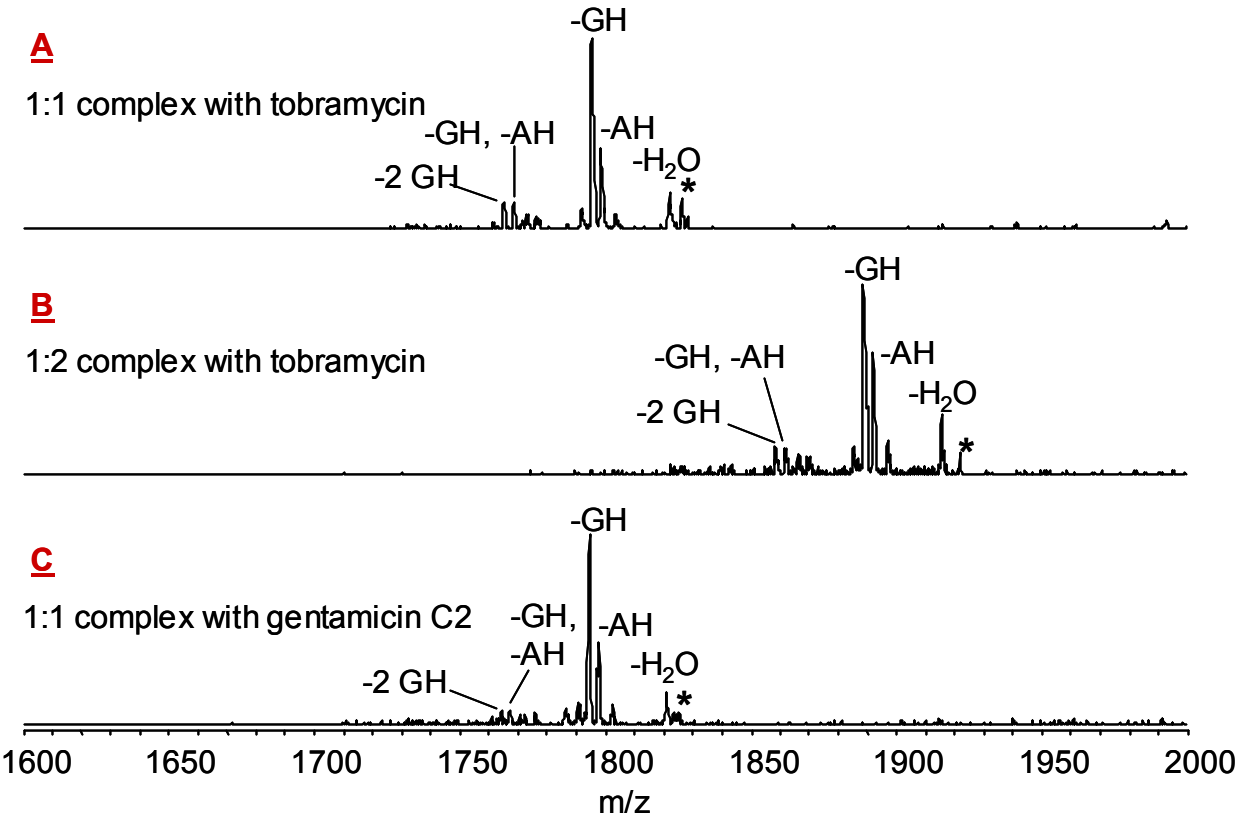


Figure 7.4: CAD of non-covalent complexes with the tobramycin aptamer (-5 charge states). (A) 1:1 and (B) 1:2 complexes with tobramycin, (C) 1:1 complex with gentamicin C2. Parent ions are marked (*); no significant fragments were observed below m/z 1600.

covalent loss of ligand (use of the –6 charge state precursor, short activation times/high activation voltages), the 1:1 aptamer:tobramycin complex still dissociated exclusively by covalent cleavage. The dissociation behavior of the 1:1 aptamer:bekanamycin complexes was essentially the same (data not shown). For both of these complexes, a network of hydrogen bonds confers high affinity binding, and because such interactions are generally stronger in the gas phase than in solution,[11,14] the energy involved in disrupting this network of interactions exceeded the energetic cost of covalent base loss.

In solution, the second tobramycin binds much less strongly (13 mM vs. 12 nM), and the 1:1 complex with gentamicin is also known to be a relatively low-affinity interaction (although a specific binding constant has not been reported). CAD mass spectra acquired for the –5 charge state precursors of both the 1:2 aptamer:tobramycin complex and the 1:1 aptamer:gentamicin C2 complex (shown in Fig. 7.4B and C) illustrate that these complexes dissociated exclusively by covalent cleavage just like the high-affinity 1:1 aptamer:tobramycin complex (Fig. 7.4A). These results were also independent of both the initial charge state of the complex and the activation time/voltage (data not shown), which indicates that electrostatic and/or hydrogen bonding interactions maintain the intermolecular contacts in these complexes as well. Because the CAD spectra reflect the nature of the gas-phase aptamer/ligand interactions rather than the solution binding constants, full scan ESI mass spectra were more reliable than the CAD mass spectra for determining the relative binding affinities of this receptor/ligand system.

7.3.2 ATP Aptamer/Ligand Interactions

The 27-residue DNA aptamer selected to bind ATP also binds the related molecules AMP and adenosine with approximately equal affinity ($K_d \sim 6$ μM).[19] An NMR study conducted at pH 6.3 showed that these molecules bind with a characteristic

1:2 aptamer:ligand stoichiometry, probably as a consequence of the high ligand load and long covalent linker used during affinity chromatography in the original selection experiments.[20] In the 1:2 aptamer:AMP complex (PDB code 1AW4), the aptamer adopts a hairpin conformation, with Watson-Crick base-pairing at the stem and an internal mismatch region that provides two sites for ligand binding.[20] Each AMP molecule intercalates between a reversed-Hoogsteen mismatch and the adenine partner of a sheared G•A mismatch, and also forms hydrogen bonds with an internal guanine residue. These recognition motifs discriminate effectively against GTP, UTP, and CMP, as well as the purine analog inosine.[19]

Several ESI mass spectra acquired for solutions containing this aptamer are shown in Figure 7.5. Solutions containing only the aptamer produced ions at m/z 1413 and 1696, corresponding to the –6 and –5 charge states; the spectral region around the –5 charge state is shown in Fig. 7.5A. A solution containing the aptamer and two equivalents of adenosine produced 1:1 and 1:2 aptamer:ligand complexes at m/z 1750 and 1803 (Fig. 7.5B). Titration experiments (not shown) demonstrated that these species do not form in stepwise fashion, which confirms that the two ligand molecules bind cooperatively in agreement with the known solution behavior.[20]

Satellite peaks at m/z 1757 and 1809 are also visible in Fig. 7.5B, and these were tentatively assigned as aptamer/K^+/ligand or aptamer/Ca^{+2}/ligand complexes. The spontaneous incorporation of adventitious metals indicates that efficient ATP aptamer/adenosine binding may require the presence of metal cations to ensure proper DNA folding. In an attempt to facilitate ligand binding, two equivalents of divalent metal salt ($MnCl_2$) were added to the aptamer/adenosine solutions, and upon electrospray ionization these solutions produced 1:1 and 1:2 aptamer:metal adducts, as well as 1:1:1 and 1:1:2 aptamer:metal:ligand complexes (Fig. 7.5D). The overall proportion of ligand-

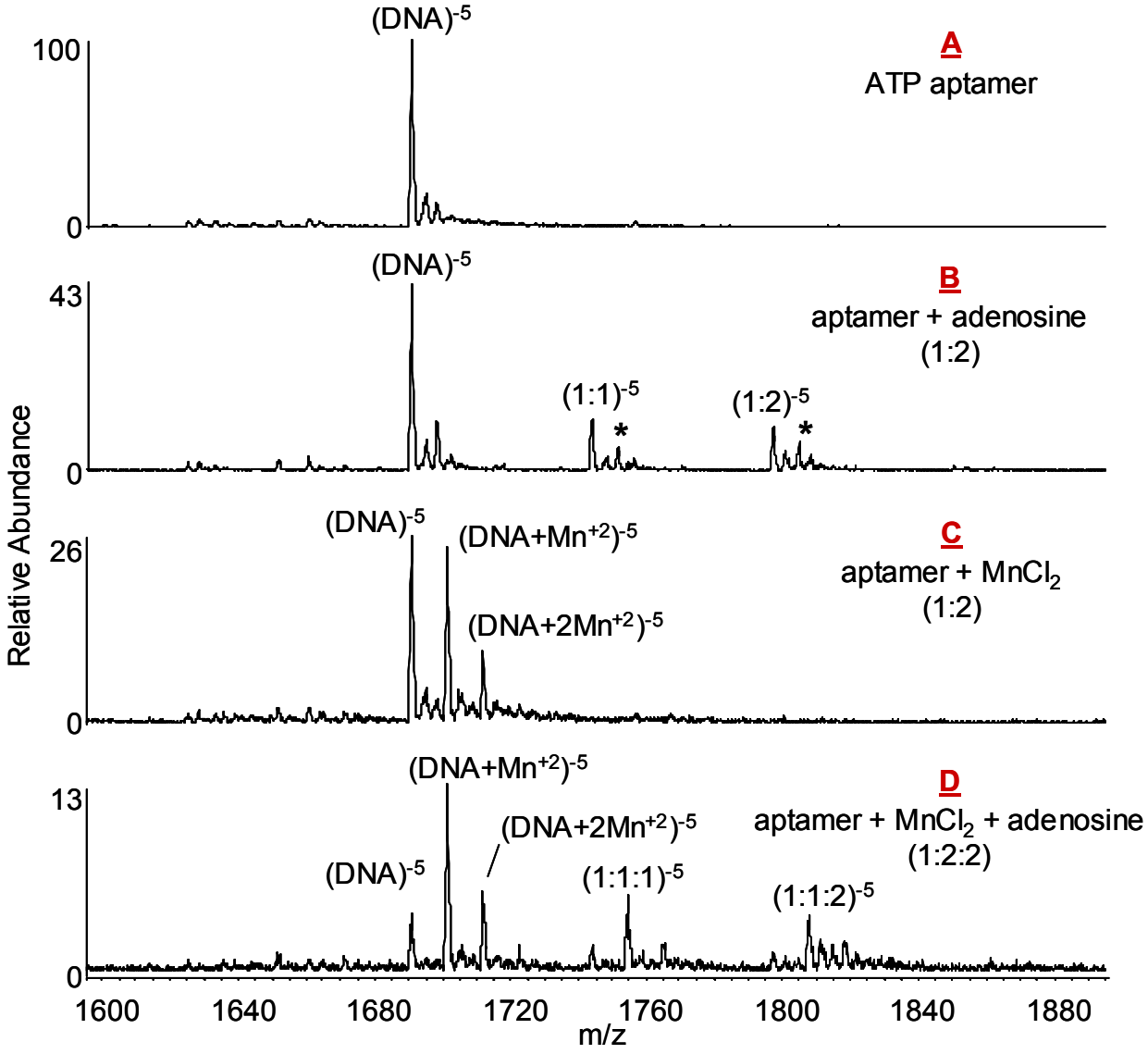


Figure 7.5: ESI spectra for samples containing the ATP aptamer (DNA). A) ATP aptamer, B) aptamer + 2 eq. adenosine, C) aptamer + 2 eq. $MnCl_2$ and D) aptamer + 2 eq. adenosine, 2 eq. $MnCl_2$. Only the spectral region containing the -5 charge states is shown. Complexes with adventitious metal ions (K^+ or Ca^{+2}) are marked *. All solutions contained 20 μM DNA and 50 mM NH_4OAc in 10% IPA/H_2O.

bound aptamer did not appear to increase with the addition of metal cations, however: the peak intensities for the aptamer:metal:ligand complexes in Fig. 7.5D were roughly the same as the peak intensities for the aptamer:ligand complexes in Fig. 7.5B. Annealing the samples prior to ESI analysis did not increase the signal intensity of the complexes, nor did the addition of excess metal or excess ligand (up to 1:16:2 and 1:2:16 aptamer:metal:adenosine ratios). A possible explanation for these results arose on examination of the mass spectrum for a solution containing aptamer and Mn^{+2} but no added ligand, shown in Fig. 7.5C. Here the $[DNA]^{-5}$ and $[DNA + Mn^{+2}]^{-5}$ ions have nearly the same signal intensity. In Fig. 7.5D, however, the $[DNA + Mn^{+2}]^{-5}$ ion is much more intense than the $[DNA]^{-5}$ ion, and this change in peak ratio must be due to the presence of adenosine in the solution. This led to the speculation that the $[DNA + Mn^{+2}]^{-5}$ ion in Fig. 7.5D might arise from both association of the free aptamer with Mn^{+2} as well as dissociation of the aptamer/metal/ligand complexes during the electrospray process. Each ligand interacts with the aptamer through the combination of π-stacking and two hydrogen bonds, and it seems intuitive that these interactions would be weakened in the gas phase because the stabilizing contribution of the hydrophobic interactions would be lost.

Some experimental evidence for the limited gas-phase stability of these complexes was found in follow-up experiments. In ESI spectra collected at a heated capillary temperature of 110°C, the 1:1:1 and 1:1:2 complexes were virtually undetectable, demonstrating that these species quite susceptible to thermal decomposition. Furthermore, in the presence of excess metal or ligand the 1:1:1 complex in Fig. 7.5D remained visible in ESI-MS spectra of aptamer/metal/adenosine solutions at approximately the same relative intensity as the 1:1:2 complex. As noted above, two adenosine molecules bind this aptamer cooperatively in solution, i.e. without a discrete

intermediate at the 1:1:1 stoichiometry. [20] It therefore seems unlikely that the 1:1:1 complex seen in Fig. 7.5D represents an intermediate in an associative process and is instead a decomposition product deriving from the 1:1:2 complex.

Additional evidence of low gas-phase stability for some ATP aptamer/ligand complexes was also seen in ligand discrimination experiments conducted with adenosine, inosine, 3'-AMP, 5'-AMP, ATP, and 5'-aminoadenosine. ESI-MS spectra acquired for six solutions containing the ATP aptamer, two equivalents of Mn^{+2} and two equivalents of one of these six ligands are shown in Fig. 7.6A-F. The apparent trend in relative binding affinity, as indicated by signal intensities for intact aptamer/metal/ligand complexes, decreases in the order 5'-aminoadenosine >> adenosine > 5'-AMP, ATP > 3'-AMP, inosine. Column binding assays were performed using ammonium acetate/$MnCl_2$ buffer as described above to provide solution data for comparison. The results, shown in Figure 7.7, indicate a solution binding trend of ATP, 5'-AMP, 5'-aminoadenosine > adenosine > 3'-AMP, inosine, which does not completely agree with the trend in signal intensity for the intact gas-phase complexes. The gas-phase trend does make sense, however, when both solution affinity *and* ligand polarity (i.e. charge character in solution at the experimental pH of 6.5) are taken into account. For example, the data in Fig. 7.7B indicate that the aptamer binds 5'-aminoadenosine efficiently in the electrospray buffer. At pH 6.5, this ligand carries a formal positive charge on the 5'-amine group. The intense signals for gas-phase aptamer/Mn^{+2}/5'-aminoadenosine complexes were therefore generated in solution by high-affinity recognition and preserved in the gas phase by dipolar and/or electrostatic interactions between the negatively charged aptamer and the positively charged ligand. (This interpretation also explains the observation why the 1:1:2 stoichiometry clearly predominates in Fig. 7.6F.) Adenosine binds well as a neutral molecule in solution (Fig. 7.7) but produces less

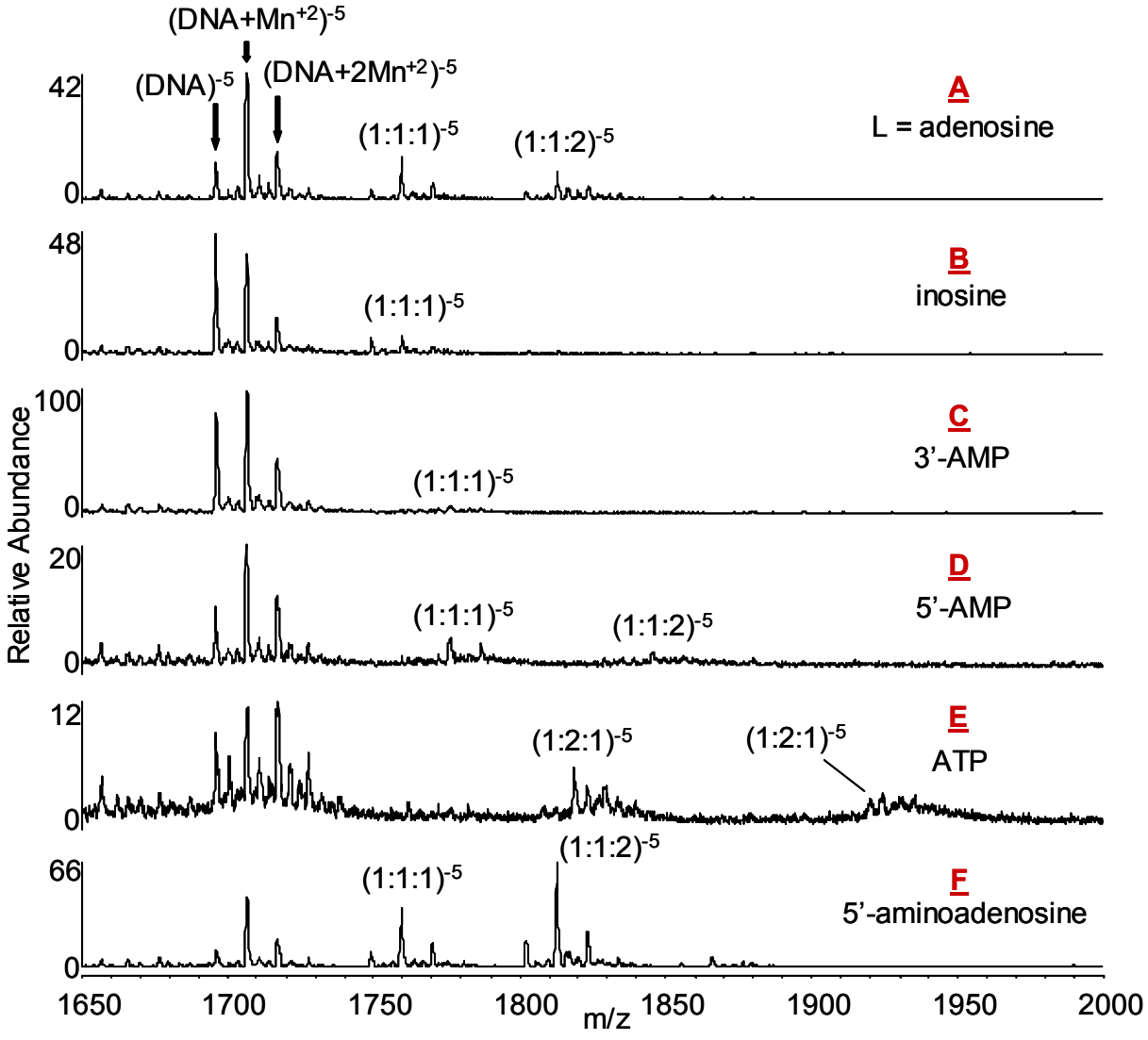


Figure 7.6: ESI spectra for the ATP aptamer (DNA), 2 eq. $MnCl_2$ and 2 eq. ligand. A) adenosine, B) inosine, C) 3'-AMP, D) 5'-AMP, E) ATP, and F) 5'-aminoadenosine. Only the spectral region containing the -5 charge states is shown. All samples contained 20 μM DNA and 50 mM NH_4OAc in 10% IPA/H_2O.

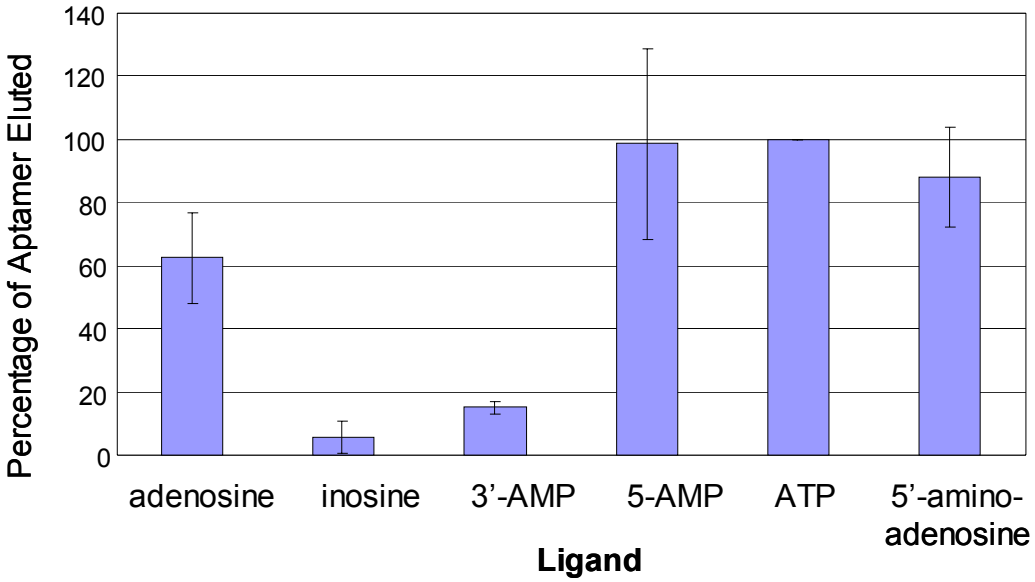


Figure 7.7: Column binding assays for ATP aptamer binding. Percentage of aptamer eluted, normalized to ATP.

intense gas-phase complexes than 5'-aminoadenosine (Fig. 7.6A vs. F) because it lacks the extra stabilization provided by the additional amine group. Both 5'-AMP and ATP carry formal negative charges in aqueous solution at pH 6.5, so while the aptamer binds these ligands efficiently in solution, Coulombic repulsion destabilizes the interactions in the gas phase and greatly suppresses signal intensity for the intact complexes (Fig. 7.6D and E). The fact that the aptamer apparently prefers to bind ATP in the presence of two Mn^{+2} cations (Fig. 7.6D) is likely another consequence of the overall charge of the ligand.

It is interesting to note, however, that the solution results roughly correlate with the trend in signal intensity patterns in the spectral region between m/z 1696 and 1730 (corresponding to the free aptamer and its Mn^{+2} complexes). For ESI spectra acquired in the presence of ligands that bind weakly in solution (inosine, Fig. 7.6B and 3'-AMP, Fig. 7.6C), the $[DNA]^{-5}$ and $[DNA + Mn^{+2}]^{-5}$ ions have nearly the same signal intensity, as can be seen in Fig. 7.5C for a solution containing no added ligand. In contrast, ESI spectra acquired for the ligands that bind strongly in solution (adenosine, 5'-AMP, ATP, and 5'-aminoadenosine, Fig. 7.6A and D-F), the $[DNA]^{-5}$ ion is less intense than the $[DNA + Mn^{+2}]^{-5}$ ions (alone and in combination with related ions at higher metal:aptamer stoichiometries). This suggests that when aptamer/ligand binding interactions are metal mediated, differences in the metal adduction patterns around signals for the free aptamer can be diagnostic of solution interactions, regardless of whether intact complexes are directly observed.

7.3.3 FMN Aptamer/Ligand Interactions

The 35-residue RNA aptamer for FMN binds its cognate ligand and the related compound 7,8-dimethylalloxazine with 500 nM affinity.[21] Structural studies of the 1:1 aptamer:FMN complex (PDB code 1FMN) have shown that the FMN aptamer,

like the ATP aptamer, adopts a hairpin-like conformation and provides a ligand binding pocket in an internal mismatch region.[22] The ligand intercalates between a G•U•A base triple and a G•G mismatch and also forms hydrogen bonds with the Hoogsteen edge of an internal adenine residue. The intermolecular interactions involved in this system, like the ATP aptamer system, therefore combine both dipole/dipole and hydrophobic interactions to achieve complex stability. A strict requirement for the presence of divalent metal cations has been noted.[22]

ESI mass spectra acquired for this aptamer are shown in Figure 7.8. The aptamer alone produced a strong signal at m/z 1889, corresponding to the –6 charge state, along with a weaker signal at m/z 1619 for the –7 charge state. The spectral region around the –6 charge state is shown in Fig. 7.8A. Initial ligand binding experiments with this aptamer were conducted using riboflavin, a close analog of FMN. ESI-MS of solutions containing the aptamer and 5 equivalents of riboflavin produced an ion at m/z 1958 that was tentatively assigned as a 1:1:1 complex between the aptamer, adventitious K^+ or Ca^{+2}, and riboflavin (Fig. 7.8B). These results were entirely consistent with the behavior of this aptamer in solution experiments.[22] Added $MnCl_2$ (5 eq.) produced the analogous aptamer/Mn^{+2}/riboflavin complex in somewhat greater signal intensity as shown in Fig. 7.8D. The $[RNA + Mn^{+2}]^{-6}/[RNA]^{-6}$ peak intensity ratio was somewhat higher when both Mn^{+2} and riboflavin were present in solution than when only Mn^{+2} was present (Fig. 7.8D vs. C). As noted in experiments with the ATP system, this shift in the distribution of free vs. metallated aptamer indicates that some degree of gas-phase instability affects noncovalent ligand binding interactions of the FMN aptamer. Such parallel results for the ATP and FMN aptamers make sense given that they bind their target ligands using similar recognition motifs (intercalation anchored by hydrogen bonds inside the binding cleft).

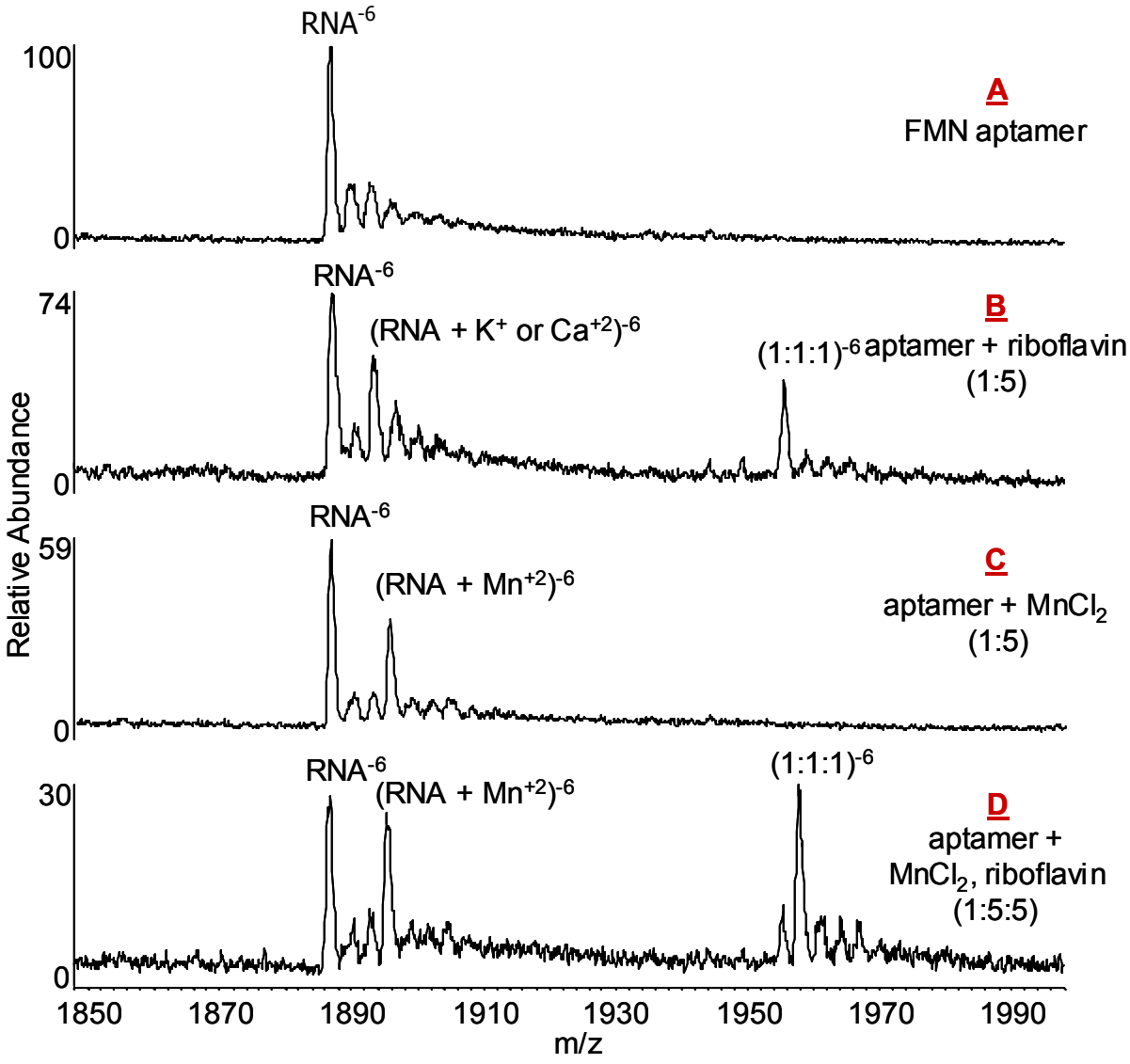


Figure 7.8 ESI spectra for solutions containing the FMN aptamer (RNA). A) FMN aptamer, B) aptamer + 5 eq. riboflavin, C) aptamer + 5 eq. $MnCl_2$, and D) aptamer + 5 eq. riboflavin, 5 eq. MnC_{l2}. Only the region containing the -6 charge statesis shown. All solutions contained 2 μM aptamer and 50 mM NH_4OAc in 30% IPA/H_2O.

ESI mass spectra for the FMN aptamer in the presence of Mn^{+2} and riboflavin, FMN, and 7,8-dimethylalloxazine are shown in Figure 7.9A-C, respectively. The solution containing FMN produced a weak signal for the 1:1:1 aptamer:Mn^{+2}:ligand complex, as well as an $[RNA + Mn^{+2}]^{-6}/[RNA]^{-6}$ peak intensity ratio that, in comparison to the data shown in Fig. 7.8C, suggests that significant decomposition of the complex has occurred in the gas phase. In light of the results obtained in the ATP aptamer system for negatively charged ligands, these results were not surprising because like 5'-AMP and ATP, FMN would have been deprotonated in the electrospray solution. The lack of stable gas-phase complexes containing 7,8-dimethylalloxazine (Fig. 7.9C) was unexpected, however, because this is a neutral ligand reported to bind the aptamer with high affinity in solution.[21] Column binding experiments using ammonium acetate buffer confirmed that riboflavin and 7,8-dimethylalloxazine bind with equal affinity in the electrospray solvent (data not shown). The $[RNA + Mn^{+2}]^{-6}/ [RNA]^{-6}$ peak intensity ratio in Fig. 7.9C is suggestive of gas-phase decomposition, but Coulombic effects must be ruled out as an underlying cause in this case. This ligand lacks the glycerol side chain of riboflavin, and it may be that this side chain interacts with the aptamer and provides additional stability for the gas phase aptamer/Mn^{+2}/riboflavin complex. (Interactions between the phosphoglycerol side chain of FMN and the RNA backbone have been postulated based on molecular modeling studies of the FMN aptamer/FMN complex.[23]) Without these additional interactions, intact complexes between the aptamer and 7,8-dimethylalloxazine may be too weak to survive in the gas phase. As for the ATP system, then, results obtained by ESI-MS do not accurately reflect solution trends for the FMN aptamer.

A CAD mass spectrum acquired for the 1:1:1 aptamer:Mn^{+2}:riboflavin complex in the –6 charge state (Fig. 7.10) showed that this precursor dissociates exclusively by loss

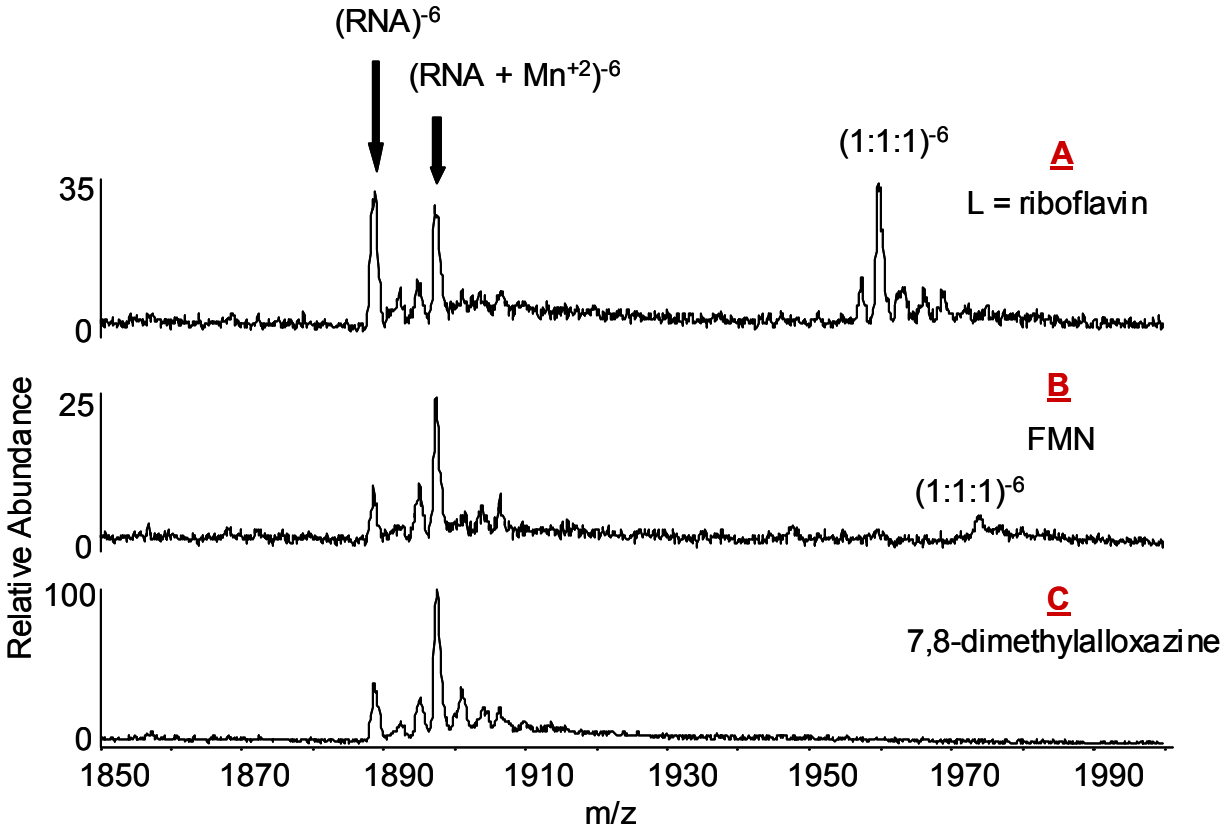


Figure 7.9: ESI spectra for the FMN aptamer + 5 eq. $MnCl_2$, 5 eq. ligand. A) Riboflavin, B) FMN, and C) 7,8-dimethylalloxazine. Only the region containing the -6 charge states is shown. All solutions contained 2 μM RNA and 50 mM NH_4OAc in 30% IPA/H_2O.

of ligand. The binding constant for this complex is approximately 500 nM and is due to the combined influence of at least two ligand-aptamer hydrogen bonds and hydrophobic interactions. The binding constant for the 1:2 tobramycin aptamer:tobramycin interaction, which is maintained by hydrogen bonding, is 13 μM, and CAD of this complex dissociated exclusively by covalent cleavage with no apparent loss of ligand. Taken in conjunction, these results indicate that during collisional activation, the type of interactions maintaining intermolecular interaction dictate the observed fragmentation pathways, and that the observed fragmentation pathways do not necessarily correlate with solution K_d.

7.3.4 Theophylline Aptamer/Ligand Interactions

The 33-residue theophylline aptamer (RNA) binds its target ligand with 300 nM affinity and, in a striking example of the ligand specificity that can be achieved with aptamers, can discriminate for theophylline vs. the closely related compound caffeine by a factor of ~ 10,000.[24] An NMR study of the 1:1 aptamer:theophylline complex (PDB code 1EHT) at pH 6.8 showed that hydrogen bonds between the ligand, one cytidine residue and one uridine residue form a base triple that interacts via π-stacking with A•C•G and U•U•A base triples above and below.[25] Divalent metal cations (Mg^{+2}, Co^{+2} or Mn^{+2}) are required to fold the RNA backbone properly.[24,26]

A series of ESI mass spectra acquired for solutions containing this RNA are shown in Figure 7.11 and illustrate that, much like the ATP and FMN systems, aptamer/ligand interactions exhibit poor gas-phase stability here as well. For the aptamer alone the dominant ion was the –6 charge state at m/z 1769 (Fig. 7.11A), and in the presence of a moderate excess of $MnCl_2$ (5 eq.) the 1:1 aptamer:Mn^{+2} ion became visible at low signal intensity (Fig. 7.11B). The solution containing theophylline (5 eq.) did not

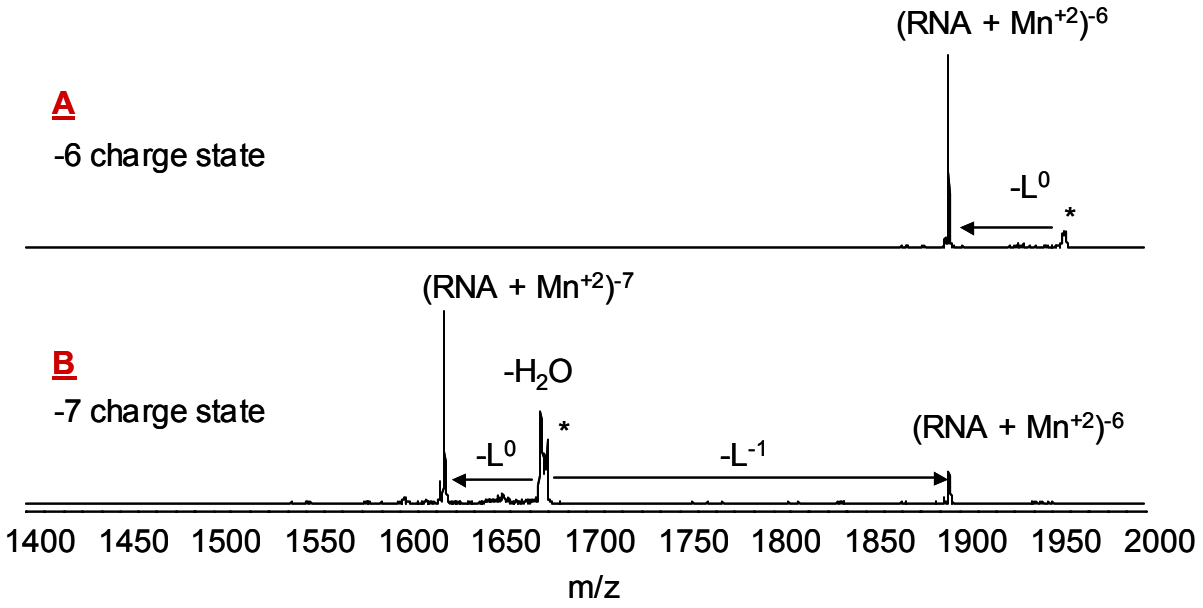


Figure 7.10: CAD of 1:1:1 complexes with the FMN aptamer, Mn^{+2}, and riboflavin in the (A) -6 and (B) -7 charge states. Parent ions are marked (*); L = riboflavin.

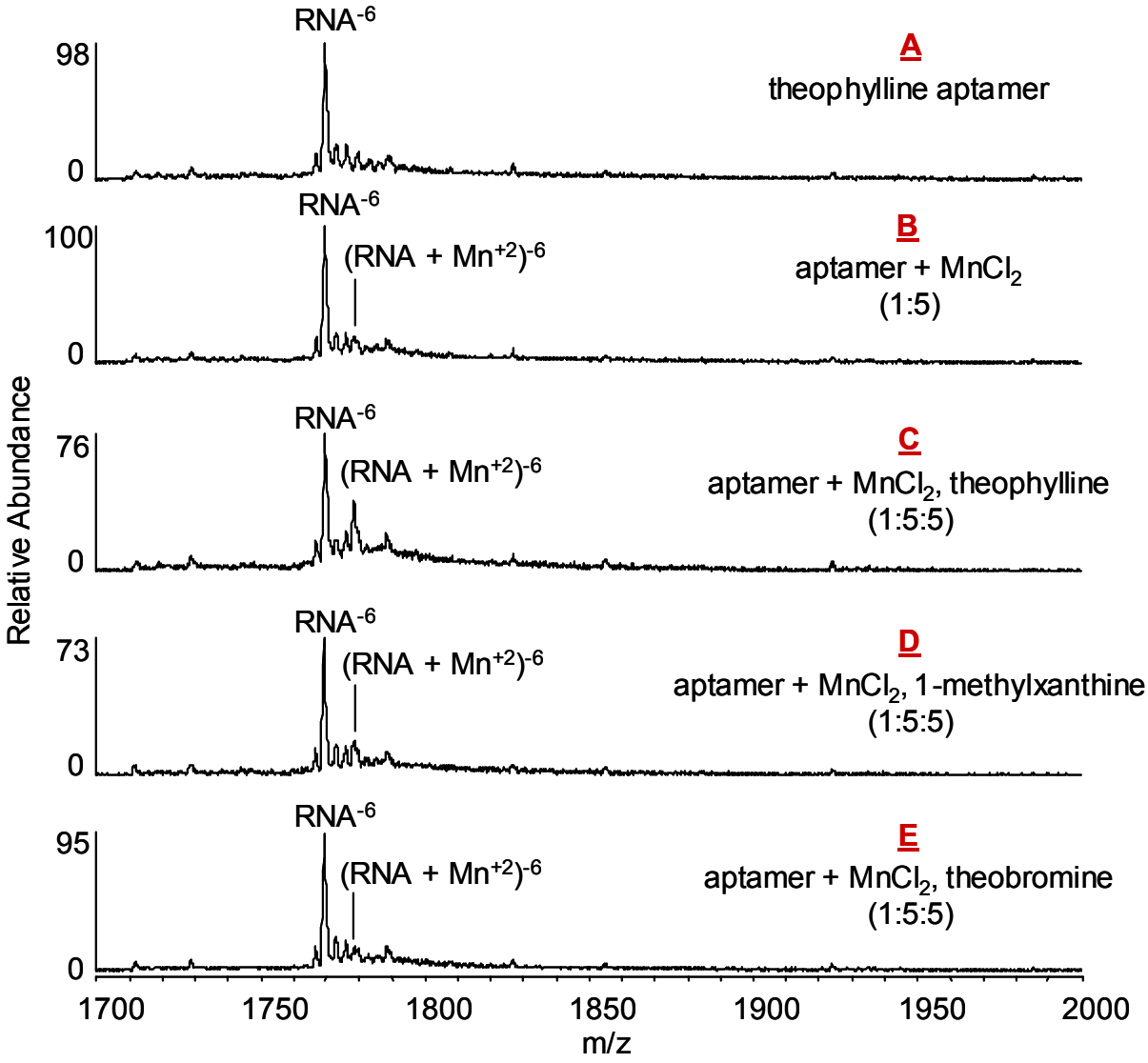


Figure 7.11: ESI spectra for solutions containing the theophylline aptamer (RNA). A) Theophylline aptamer, B) aptamer + 5 eq. $MnCl_2$, C) aptamer + 5 eq. $MnCl_2$, 5 eq. theophylline, D) aptamer + 5 eq. $MnCl_2$, 5 eq. 1-methylxanthine, E) aptamer + 5 eq. $MnCl_2$, 5 eq. theobromine. Only the region containing the -6 charge states is shown. All solutions contained 2 μM aptamer and 50 mM NH_4OAc in 30% IPA/H_2O.

produce intact aptamer/ligand complexes, however (Fig. 7.11C). In this system column binding experiments were not conducted to verify that aptamer/ligand binding occurs in the electrospray buffer. However, the intensity of the $[RNA + Mn^{+2}]^{-6}$ ion is greater in Fig. 7.11C than in Fig. 7.11B, which may be diagnostic of solution interactions as noted in the interpretation of MS data acquired for the ATP and FMN aptamers. Furthermore, the relative intensity of the $[RNA + Mn^{+2}]^{-6}$ ion decreases for solutions containing theophylline, 1-methylxanthine, and theobromine (Figs. 7.11C-E), and this trend correlates exactly with the reported solution binding affinities of these compounds.[24] These observations support the notion that the gas-phase stability of these aptamer/ligand complexes is compromised because the π-stacking forces that contribute so much stability in solution lose potency *in vacuo*.

7.4 CONCLUSIONS

Electrospray ionization mass spectrometry (ESI-MS) has been used to assess molecular recognition of small molecule ligands by nucleic acid aptamers for tobramycin, ATP, FMN, and theophylline. The overall results indicate that the gas-phase behavior of aptamer/ligand complexes correlates with solution behavior in only some respects. The binding stoichiometry and relative binding affinities determined by ESI-MS for the tobramycin aptamer, which interacts with its cognate ligand through hydrogen bonding interactions, accurately reflected the known solution behavior. Results for the ATP, FMN, and theophylline aptamers, all of which recognize ligands via both hydrogen bonding and hydrophobic interactions, were more problematic. For these systems, the metal dependence of ligand binding in solution was accurately reflected in the mass spectral data. However, Coulombic effects deriving from ligand polarity/charge character stabilized some high-affinity aptamer/ligand complexes and destabilized others,

and complexes with a greater number of intermolecular hydrogen bonds exhibited greater gas-phase stability even in cases when solution binding affinities were equivalent. Both factors precluded accurate determination of relative binding affinities from ESI-MS spectra. This indicates that complementary techniques may be required to fully assess aptamer molecular recognition properties.

7.5 REFERENCES

(1) Ellington, A. D. *Curr. Biol.* **1994**, *4*, 427-429.

(2) Conrad, R. C.; Bruck, F. M.; Bell, S.; Ellington, A. D. *RNA Prot. Interact.* **1998**, 285-325.

(3) Famulok, M.; Mayer, G.; Blind, M. *Acc. Chem. Res.* **2000**, *33*, 591-599.

(4) Hermann, T.; Patel, D. J. *Science* **2000**, *287*, 820-825.

(5) Hesselberth, J.; Robertson, M. P.; Jhaveri, S.; Ellington, A. D. *Rev. Mol. Biotechnol.* **2000**, *74*, 15-25.

(6) Osborne, S. E.; Ellington, A. D. *Chem. Rev.* **1997**, *97*, 349-370.

(7) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611-647.

(8) Burgstaller, P.; Jenne, A.; Blind, M. *Curr. Opin. Drug Disc. Dev.* **2002**, *5*, 690-700.

(9) Sullenger, B. A.; Gilboa, E. *Nature* **2002**, *418*, 252-258.

(10) Brodbelt, J. S. *Int. J. Mass Spectrom.* **2000**, *200*, 57-69.

(11) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.

(12) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(13) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.

(14) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(15) Cassiday, L. A.; Lebruska, L. L.; Benson, L. M.; Naylor, S.; Owen, W. G.; Maher, L. J., 3rd *Anal. Biochem.* **2002**, *306*, 290-297.

(16) Cavanagh, J.; Benson, L. M.; Thompson, R.; Naylor, S. *Anal. Chem.* **2003**, *75*, 3281-3286.

(17) Wang, Y.; Rando, R. R. *Chem. Biol.* **1995**, *2*, 281-290.

(18) Jiang, L.; Patel, D. J. *Nat. Struct. Biol.* **1998**, *5*, 769-774.

(19) Huizenga, D. E.; Szostak, J. W. *Biochemistry* **1995**, *34*, 656-665.

(20) Lin, C. H.; Patel, D. J. *Chem. Biol.* **1997**, *4*, 817-832.

(21) Burgstaller, P.; Famulok, M. *Angew. Chem.* **1994**, *106*, 1163-1166.

(22) Fan, P.; Suri, A. K.; Fiala, R.; Live, D.; Patel, D. J. *J. Mol. Biol.* **1996**, *258*, 480-500.

(23) Schneider, C.; Suhnel, J. *Biopolymers* **1999**, *50*, 287-302.

(24) Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B. *Science* **1994**, *263*, 1425-1429.

(25) Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J.-P.; Pardi, A. *Nat. Struct. Biol.* **1997**, *4*, 644-649.

(26) Zimmermann, G. R.; Wick, C. L.; Shields, T. P.; Jenison, R. D.; Pardi, A. *RNA* **2000**, *6*, 659-667.

(27) In.

(28) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(29) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(30) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

(31) Keller, K. M.; Oehlers, L.; Brodbelt, J. S. **2004**, manuscript in preparation.

Chapter 8: Conclusion

Over the last two decades mass spectrometry has gained extraordinary prominence in the study of biological macromolecules. The initial (and probably most obvious) applications of MS in protein and nucleic acid analysis involved analyte identification through sequence determination, and although MS already competes well with alternative sequencing techniques such efforts still attract sustained interest. Chapters 3 and 4 of this work describe experiments conducted to characterize new or under-utilized fragmentation methods (MSAD and IRMPD) for biopolymer sequencing. These studies illustrate that different fragmentation methods can produce different information because of key details about the ion activation process, but they also show that alternative dissociation methods can nonetheless be effective and sometimes advantageous for protein and nucleic acid sequence determination. A thorough evaluation of such methods makes it possible to fully exploit the range of options available for MS sequencing experiments.

Mass spectrometry can also be used probe the higher order structure and function of biopolymers by examination of their non-covalent associations. Chapters 5, 6 and 7 presented results obtained from mass spectrometric studies of nucleic acid complexes, and the results make it clear that the specific intermolecular interactions can be reflected in mass spectrometric data. Substantial questions remain about the use of gas-phase measurements to draw conclusions about solution-phase structure, however, so a thorough understanding of the impact of different experimental variables is of paramount importance. Chapters 5 and 6 illustrated that initial charge state can strongly influence

gas-phase fragmentation patterns, which highlights choice of precursor ion as a major consideration in designing dissociation experiments to study biomolecular interactions. In Chapter 7, Coulombic effects emerged as a critical factor controlling which species survive and can be studied in the gas phase, a result that suggests that MS is not universally applicable to the study of all non-covalent complexes.

Collectively, these studies supply new information on the gas-phase stability and dissociation tendencies of proteins, nucleic acids, and nucleic acid/small molecule complexes, and the results point the way to improved strategies for the identification and characterization of biopolymers that fully exploit the rapid, sensitive nature of mass spectrometric analysis. Yet they also point to particular problems and limitations that should be addressed in future MS research. At present mass spectrometry cannot provide routine *de novo* sequencing of large biopolymers, and new dissociation techniques and analytical strategies will be required to achieve this goal. In addition, while even the most rudimentary biochemistry textbooks stress the importance of tertiary and quaternary structure on the function of proteins and nucleic acids in solution, very little is known about the conformation of these molecules in the gas phase. Further growth and refinement in the use of ion mobility and other techniques that evaluate conformation will produce new information on gas-phase structure, which should not only provide new insight on folding patterns and dynamics observed in solution but should also provide a more detailed framework for analysis of non-covalent interactions that are preserved in the gas phase. Continued development in areas such as these should enhance the overall scope and utility of mass spectrometry in biopolymer analysis.

Bibliography

(1) Voet, D.; Voet, J. G. *Biochemistry*; 2nd ed.; John Wiley & Sons:, 1995.

(2) McLafferty, F. W. *Science* **1981**, *214*, 280-287.

(3) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(4) Beckey, H. D. *Research/Development* **1969**, *20*, 26-29.

(5) Beckey, H. D. *Principles of Field Ionization and Field Desorption Mass Spectrometry*; Pergamon: Oxford, 1977.

(6) Torgerson, D. F.; Skowronski, R. P.; Macfarlane, R. D. *Biochem. Biophys. Res. Commun.* **1974**, *60*, 616-621.

(7) Macfarlane, R. D.; Torgerson, D. F. *Science* **1976**, *191*, 920-925.

(8) Sundqvist, B.; Macfarlane, R. D. *Mass Spectrom. Rev.* **1985**, *4*, 421-460.

(9) Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tyler, A. N. *Nature* **1981**, *293*, 270-275.

(10) Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tyler, A. N. *J. Chem. Soc. Chem. Commun.* **1981**, 325-327.

(11) Barber, M.; Bordoli, R. S.; Elliott, G. J.; Sedgwick, R. D.; Tyler, A. N. *Anal. Chem.* **1982**, *54*, 645A-646A, 649A-650A, 653A, 655A, 657A.

(12) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240-2249.

(13) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287-2294.

(14) Yamashita, M.; Fenn, J. B. *J. Chem. Phys.* **1984**, *88*, 4451-4459.

(15) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.

(16) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

(17) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yohida, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.

(18) Morris, H. R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R. S.; Hoyes, J.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 889-896.

(19) Qian, M. G.; Lubman, D. M. *Anal. Chem.* **1995**, *67*, 234A-236A, 238A, 240A-232A.

(20) Qian, M. G.; Lubman, D. M. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1911-1920.

(21) Hager, J. W. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 512-526.

(22) Schwartz, J. C.; Senko, M. W.; Syka, J. E. P. *J. Am. Soc.Mass Spectrom.* **2002**, *13*, 659-669.

(23) Bruins, A. P.; Covey, T. R.; Henion, J. D. *Anal. Chem.* **1987**, *59*, 2642-2646.

(24) Olivares, J. A.; Nguyen, N. T.; Yonker, C. R.; Smith, R. D. *Anal. Chem.* **1987**, *59*, 1230-1232.

(25) Smith, R. D.; Olivares, J. A.; Nguyen, N. T.; Udseth, H. R. *Anal. Chem.* **1988**, *60*, 436-441.

(26) Yang, L.; Lee, C. S.; Hofstadler, S. A.; Pasa-Tolic, L.; Smith, R. D. *Anal. Chem.* **1998**, *70*, 3235-3241.

(27) Chaurand, P.; Schwartz, S. A.; Caprioli, R. M. *J. Proteome Res.* **2004**, *3*, 245-252.

(28) Biemann, K.; Vetter, W. *Biochem. Biophys. Res. Commun.***1960**, *3*, 578-584.

(29) Papayannopoulos, I. A. *Mass Spectrom. Rev.* **1995**, *14*, 49-73.

(30) Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *J. Mass Spectrom.* **1996**, *31*, 339-350.

(31) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.

(32) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.

(33) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

(34) Kruger, N. A.; Zubarev, R. A.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom.* **1999**, *185*, 787-793.

(35) Sze, S. K.; Ge, Y.; Oh, H.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U S A* **2002**, *99*, 1774-1779.

(36) Hunt, D. F.; Yates, J. R., III; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U S A* **1986**, *83*, 6233-6237.

(37) Yates, J. R. *J. Mass Spectrom.* **1998**, *33*, 1-19.

(38) Reid, G. E.; McLuckey, S. A. *J. Mass Spectrom.* **2002**, *37*, 663-675.

(39) Kelleher, N. L. *Anal. Chem.* **2004**, *76*, 196A-203A.

(40) Meng, F.; Cargile, B. J.; Miller, L. M.; Forbes, A. J.; Johnson, J. R.; Kelleher, N. L. *Nat. Biotechnol.* **2001**, *19*, 952-957.

(41) Lipton, M. S.; Pasa-Tolic, L.; Anderson, G. A.; Anderson, D. J.; Auberry, D. L.; Battista, J. R.; Daly, M. J.; Fredrickson, J.; Hixson, K. K.; Kostandarithes, H.; Masselon, C.; Markillie, L. M.; Moore, R. J.; Romine, M. F.; Shen, Y.; Stritmatter, E.; Tolic, N.; Udseth, H. R.; Venkateswaran, A.; Wong, K.-K.; Zhao, R.; Smith, R. D. *Proc. Natl. Acad. Sci. U S A* **2002**, *99*, 11049-11054.

(42) Meng, F.; Cargile, B. J.; Patrie, S. M.; Johnson, J. R.; McLoughlin, S. M.; Kelleher, N. L. *Anal. Chem.* **2002**, *74*, 2923-2929.

(43) VerBerkmoes, N. C.; Bundy, J. L.; Hauser, L.; Asano, K. G.; Razumovskaya, J.; Larimer, F.; Hettich, R. L.; Stephenson, J. L., Jr. *J. Proteome Res.* **2002**, *1*, 239-252.

(44) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.

(45) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(46) Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1991**, *113*, 8534-8535.

(47) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294-6296.

(48) Ganem, B.; Li, Y. T.; Henion, J. D. *Tetrahedron Lett.* **1993**, *34*, 1445-1448.

(49) Light-Wahl, K. J.; Springer, D. L.; Winger, B. E.; Edmonds, C. G.; Camp, D. G., II; Thrall, B. D.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, *115*, 803-804.

(50) Doktycz, M. J.; Habibi-Goudarzi, S.; McLuckey, S. A. *Anal. Chem.* **1994**, *66*, 3416-3422.

(51) Sannes-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264-271.

(52) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(53) Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464-467.

(54) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(55) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(56) Reyzer, M. L.; Brodbelt, J. S.; Kerwin, S. M.; Kumar, D. *Nucleic Acids Res.* **2001**, *29*, E103-103.

(57) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(58) Goodlett, D. R.; Camp, D. G., II; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, *22*, 181-183.

(59) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.

(60) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.

(61) Loo, J. A.; Hu, P.; McConnell, P.; Mueller, W. T.; Sawyer, T. K.; Thanabal, V. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 234-243.

(62) Hofstadler, S. A.; Sannes-Lowery, K. A.; Crooke, S. T.; Ecker, D. J.; Sasmor, H.; Manalili, S.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 3436-3440.

(63) Sannes-Lowery, K. A.; Drader, J. J.; Griffey, R. H.; Hofstadler, S. A. *Proc. SPIE-Int. Soc. Opt. Eng.* **2001**, *4264*, 27-36.

(64) Cummins, L. L.; Chen, S.; Blyn, L. B.; Sannes-Lowery, K. A.; Drader, J. J.; Griffey, R. H.; Hofstadler, S. A. *J. Natural Prod.* **2003**, *66*, 1186-1190.

(65) Lim, H.-K.; Hsieh, Y. L.; Ganem, B.; Henion, J. *J. Mass Spectrom.* **1995**, *30*, 708-714.

(66) Greig, M. J.; Gaus, H.; Cummins, L. L.; Sasmor, H.; Griffey, R. H. *J. Am. Chem. Soc.* **1995**, *117*, 10765-10766.

(67) Kraunsoe, J. A. E.; Aplin, R. T.; Green, B.; Lowe, G. *FEBS Lett.* **1996**, *396*, 108-112.

(68) Rogniaux, H.; Van Dorsselaer, A.; Barth, P.; Biellmann, J. F.; Barbanton, J.; Van Zandt, M.; Chevrier, B.; Howard, E.; Mitschler, A.; Potier, N.; Urzhumtseva, L.; Moras, D.; Podjarny, A. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 635-647.

(69) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, *120*, 9605-9613.

(70) Gabelica, V.; De Pauw, E. *Int. J. Mass Spectrom.* **2002**, *219*, 151-159.

(71) Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. *J. Am. Chem. Soc.* **1996**, *118*, 8646-8653.

(72) Li, Y. T.; Hsieh, Y. L.; Henion, J. D.; Ocain, T. D.; Schiehser, G. A.; Ganem, B. *J. Am. Chem. Soc.* **1994**, *116*, 7487-7493.

(73) Wu, Q.; Gao, J.; Joseph-McCarthy, D.; Sigal, G. B.; Bruce, J. E.; Whitesides, G. M.; Smith, R. D. *J. Am. Chem. Soc.* **1997**, *119*, 1157-1158.

(1) Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. *Anal. Chim. Acta* **1991**, *246*, 211-225.

(2) Sannes-Lowery, K.; Griffey, R. H.; Kruppa, G. H.; Speir, J. P.; Hofstadler, S. A. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1957-1961.

(3) Dunbar, R. C. *Int. J. Mass Spectrom.* **2000**, *200*, 571-589.

(4) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, *88*, 4451-4459.

(5) Cole, R. B., Ed. *Electrospray Ionization Mass Spectrometry*; 1st ed.; Wiley-Interscience: New York, 1997.

(6) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.

(7) March, R. E. *J. Mass Spectrom.* **1997**, *32*, 351-369.

(8) Hakansson, K.; Axelsson, J.; Palmblad, M.; Hakansson, P. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 210-217.

(9) Sannes-Lowery, K. A.; Hofstadler, S. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1-9.

(1) Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001**, *101*, 269-295.

(2) Patterson, S. D. *Anal. Biochem.* **1994**, *221*, 1-15.

(3) Jensen, P. K.; Pasa-Tolic, L.; Peden, K. K.; Martinovic, S.; Lipton, M. S.; Anderson, G. A.; Tolic, N.; Wong, K. K.; Smith, R. D. *Electrophoresis* **2000**, *21*, 1372-1380.

(4) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R., 3rd *Nat. Biotechnol.* **1999**, *17*, 676-682.

(5) Hunt, D. F.; Yates, J. R., III; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U S A* **1986**, *83*, 6233-6237.

(6) Kassel, D. B.; Musselman, B. D.; Smith, J. A. *Anal. Chem.* **1991**, *63*, 1091-1097.

(7) Wilm, M.; Shevchenko, A.; Houthaeve, T.; Breit, S.; Schweigerer, L.; Fotsis, T.; Mann, M. *Nature* **1996**, *379*, 466-469.

(8) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 806-812.

(9) Reid, G. E.; McLuckey, S. A. *J. Mass Spectrom.* **2002**, *37*, 663-675.

(10) VerBerkmoes, N. C.; Bundy, J. L.; Hauser, L.; Asano, K. G.; Razumovskaya, J.; Larimer, F.; Hettich, R. L.; Stephenson, J. L., Jr. *J. Proteome Res.* **2002**, *1*, 239-252.

(11) Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. *Anal. Chim. Acta* **1991**, *246*, 211-225.

(12) Senko, M. W.; Speir, J. P.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2801-2808.

(13) Senko, M. W.; Beu, S. C.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 415-418.

(14) Smith, R. D.; Loo, J. A.; Barinaga, C. J.; Edmonds, C. G.; Udseth, H. R. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 53-65.

(15) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

(16) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

(17) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563-573.

(18) McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 245-249.

(19) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U S A* **2000**, *97*, 10313-10317.

(20) Mann, M.; Wilm, M. *Anal. Chem.* **1994**, *66*, 4390-4399.

(21) Mortz, E.; O'Connor, P. B.; Roepstorff, P.; Kelleher, N. L.; Wood, T. D.; McLafferty, F. W.; Mann, M. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8264-8267.

(22) Cargile, B. J.; McLuckey, S. A.; Stephenson, J. L. *Anal. Chem.* **2001**, *73*, 1277-1285.

(23) Demirev, P. A.; Ramirez, J.; Fenselau, C. *Anal. Chem.* **2001**, *73*, 5725-5731.

(24) Nemeth-Cawley, J. F.; Rouse, J. C. *J. Mass Spectrom.* **2002**, *37*, 270-282.

(25) Meng, F.; Cargile, B. J.; Miller, L. M.; Forbes, A. J.; Johnson, J. R.; Kelleher, N. L. *Nat. Biotechnol.* **2001**, *19*, 952-957.

(26) Reid, G. E.; Shang, H.; Hogan, J. M.; Lee, G. U.; McLuckey, S. A. *J. Am. Chem. Soc.* **2002**, *124*, 7353-7362.

(27) In https://prosightptm.scs.uiuc.edu; The University of Illinois at Urbana-Champaign.

(28) Van Berkel, G. J.; Kennel, S. J.; Doktycz, M. J.; Ford, M. J.; Sanchez, A. D.; Quirke, J. M. E. In *Proceedings of the 51st ASMS Conference on Mass Spectrometry and Allied Topics*: Montreal, Canada, 2003.

(29) Li, W.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Anal. Chem.* **1999**, *71*, 4397-4402.

(30) Sannes-Lowery, K.; Griffey, R. H.; Kruppa, G. H.; Speir, J. P.; Hofstadler, S. A. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1957-1961.

(31) Sannes-Lowery, K. A.; Hofstadler, S. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1-9.

(32) Hakansson, K.; Axelsson, J.; Palmblad, M.; Hakansson, P. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 210-217.

(33) Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal. Chem.* **2001**, *73*, 3274-3281.

(34) Newton, K. A.; Chrisman, P. A.; Reid, G. E.; Wells, J. M.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2001**, *212*, 359-376.

(35) Hogan, J. M.; McLuckey, S. A. *J. Mass Spectrom.* **2003**, *38*, 245-256.

(36) Belov, M. E.; Gorshkov, M. V.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1312-1319.

(37) In http://us.expasy.org/tools/tagident.html; Swiss Institute of Bioinformatics/North Carolina Supercomputing Center.

(38) Tolmachev, A. V.; Udseth, H. R.; Smith, R. D. *Int. J. Mass Spectrom.* **2003**, *222*, 155-174.

(1) Colorado, A.; Shen, J. X.; Brodbelt, J. *Anal. Chem.* **1996**, *68*, 4033-4043.

(2) Goolsby, B. J.; Brodbelt, J. S. *J. Mass Spectrom.* **1998**, *33*, 705-712.

(3) Vartanian, V. H.; Goolsby, B.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1089-1098.

(4) Goolsby, B. J.; Brodbelt, J. S. *J. Mass Spectrom.* **2000**, *35*, 1011-1024.

(5) Crowe, M. C.; Brodbelt, J. S.; Goolsby, B. J.; Hergenrother, P. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 630-649.

(6) Stephenson, J. L., Jr.; Booth, M. M.; Shalosky, J. A.; Eyler, J. R.; Yost, R. A. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 886-893.

(7) Stephenson, J. L., Jr.; Booth, M. M.; Boue, S. M.; Eyler, J. R.; Yost, R. A. *ACS Symp. Ser.* **1996**, *619*, 512-564.

(8) Boue, S. M.; Stephenson, J. L., Jr.; Yost, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1391-1397.

(9) Payne, A. H.; Glish, G. L. *Anal. Chem.* **2001**, *73*, 3542-3548.

(10) Hashimoto, Y.; Hasegawa, H.; Yoshinari, K.; Waki, I. *Anal. Chem.* **2003**, *75*, 420-425.

(11) Drader, J. J.; Hannis, J. C.; Hofstadler, S. A. *Anal. Chem.* **2003**, *75*, 3669-3674.

(12) Jockusch, R. A.; Paech, K.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 3188-3196.

(13) Masselon, C.; Anderson, G. A.; Harkewicz, R.; Bruce, J. E.; Pasa-Tolic, L.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 1918-1924.

(14) Flora, J. W.; Muddiman, D. C. *Anal. Chem.* **2001**, *73*, 3305-3311.

(15) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.

(16) Li, L.; Masselon, C. D.; Anderson, G. A.; Pasa-Tolic, L.; Lee, S. W.; Shen, Y.; Zhao, R.; Lipton, M. S.; Conrads, T. P.; Tolic, N.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 3312-3322.

(17) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

(18) Mortz, E.; O'Connor, P. B.; Roepstorff, P.; Kelleher, N. L.; Wood, T. D.; McLafferty, F. W.; Mann, M. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8264-8267.

(19) Li, W.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Anal. Chem.* **1999**, *71*, 4397-4402.

(20) Meng, F.; Cargile, B. J.; Miller, L. M.; Forbes, A. J.; Johnson, J. R.; Kelleher, N. L. *Nat. Biotechnol.* **2001**, *19*, 952-957.

(21) Meng, F.; Cargile, B. J.; Patrie, S. M.; Johnson, J. R.; McLoughlin, S. M.; Kelleher, N. L. *Anal. Chem.* **2002**, *74*, 2923-2929.

(22) Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893-4897.

(23) Little, D. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **1995**, *117*, 6783-6784.

(24) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **1996**, *118*, 9352-9359.

(25) Hofstadler, S. A.; Griffey, R. H.; Pasa-Tolic, L.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1400-1404.

(26) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 2067-2070.

(27) Hofstadler, S. A.; Sannes-Lowery, K. A.; Griffey, R. H. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 945-951.

(28) Sannes-Lowery, K. A.; Hofstadler, S. A. *J. Am. Soc.Mass Spectrom.* **2003**, *14*, 825-833.

(29) Crowe, M. C.; Brodbelt, J. S. In *15th Sanibel Conference on Mass Spectrometry*: Sanibel, FL, 2003.

(30) Black, D. M.; Stephens, J. D.; Glish, G. L. In *Proceedings of the 51st ASMS Conference on Mass Spectrometry and Allied Topics*: Montreal, Canada, 2003.

(31) Keller, K. M.; Ostrander, C. M.; Fannin, S. T.; Brodbelt, J. S. In *Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics*: Chicago, IL, 2001.

(32) Reid, G. E.; Mitchell Wells, J.; Badman, E. R.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2003**, *222*, 243-258.

(33) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.

(34) McLuckey, S. A.; Habibi-Goudarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085-12095.

(35) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Processes* **1994**, *137*, 121-149.

(36) Barry, J. P.; Vouros, P.; Van Schepdael, A.; Law, S.-J. *J. Mass Spectrom.* **1995**, *30*, 993-1006.

(37) Wang, Z.; Wan, K. X.; Ramanathan, R.; Taylor, J. S.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 683-691.

(38) Wan, K. X.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 580-589.

(39) Wan, K. X.; Gross, J.; Hillenkamp, F.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 193-205.

(40) Premstaller, A.; Huber, C. G. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1053-1060.

(41) McLuckey, S. A.; Vaidyanathan, G.; Habibi-Goudarzi, S. *J. Mass Spectrom.* **1995**, *30*, 1222-1229.

(42) Sannes-Lowery, K. A.; Mack, D. P.; Hu, P.; Mei, H.-Y.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 90-95.

(43) Ni, J.; Mathews, M. A. A.; McCloskey, J. A. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 535-540.

(44) Wang, P.; Bartlett, M. G.; Martin, L. B. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 846-856.

(45) Weimann, A.; Iannitti-Tito, P.; Sheil, M. M. *Int. J. Mass Spectrom.* **2000**, *194*, 269-288.

(46) Vrkic, A. K.; O'Hair, R. A. J.; Foote, S.; Reid, G. E. *Int. J. Mass Spectrom.* **2000**, *194*, 145-164.

(47) Boschenok, J.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 144-149.

(48) Hakansson, K.; Hudgins, R. R.; Marshall, A. G.; O'Hair, R. A. J. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 23-41.

(1) Hud, N. V.; Polak, M. *Curr. Opin. Struct. Biol.* **2001**, *11*, 293-301.

(2) McFail-Isom, L.; Sines, C. C.; Williams, L. D. *Curr. Opin. Struct. Biol.* **1999**, *9*, 298-304.

(3) Frank-Kamenetskii, M. D.; Mirkin, S. M. *Annu. Rev. Biochem.* **1995**, *64*, 65-95.

(4) Gilbert, D. E.; Feigon, J. *Curr. Opin. Struct. Biol.* **1999**, *9*, 305-314.

(5) Chowdhury, S.; Bansal, M. *J. Biomol. Struct. Dyn.* **2000**, *18*, 11-28.

(6) Hardin, C. C.; Perry, A. G.; White, K. *Biopolymers* **2000**, *56*, 147-194.

(7) Draper, D. E.; Shiman, R. *J. Mol. Biol.* **2000**, *302*, 79-91.

(8) Draper, D. E. *RNA* **2004**, *10*, 335-343.

(9) Chakrabarti, S.; Bhattacharyya, D.; Dasgupta, D. *Biopolymers* **2000**, *56*, 85-95.

(10) Fan, J. Y.; Sun, D.; Yu, H.; Kerwin, S. M.; Hurley, L. H. *J. Med. Chem.* **1995**, *38*, 408-424.

(11) Zeng, Q.; Kwok, Y.; Kerwin, S. M.; Mangold, G.; Hurley, L. H. *J. Med. Chem.* **1998**, *41*, 4273-4278.

(12) Ni, J.; Pomerantz, S. C.; Rozenski, J.; Zhang, Y.; McCloskey, J. A. *Anal. Chem.* **1996**, *68*, 1989-1999.

(13) Nordhoff, E.; Kirpekar, F.; Roepstorff, P. *Mass Spectrom. Rev.* **1996**, *15*, 69-138.

(14) Limbach, P. A. *Mass Spectrom. Rev.* **1997**, *15*, 297-336.

(15) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.

(16) Gonnet, F.; Kocher, F.; Blais, J.; Bolbach, G.; Chottard, J.; Tabet, J. C. *Inorg. Chem.* **1996**, *35*, 1653-1658.

(17) Lowe, G.; McCloskey, J. A.; Ni, J.; Vilaivan, T. *Bioinorg. Med. Chem.* **1996**, *4*, 1007-1013.

(18) Troujman, H.; Chottard, J. C. *Anal. Biochem.* **1997**, *252*, 177-185.

(19) Xu, N.; Pasa-Tolic, L.; Smith, R. D.; Ni, S.; Thrall, B. D. *Anal. Biochem.* **1999**, *272*, 26-33.

(20) Iannitti-Tito, P.; Weimann, A.; Wickham, G.; Sheil, M. M. *Analyst* **2000**, *125*, 627-634.

(21) Kloster, M. B. G.; Hannis, J. C.; Muddiman, D. C.; Farrell, N. *Biochemistry* **1999**, *38*, 14731-14737.

(22) Carte, N.; Legendre, F.; Leize, E.; Potier, N.; Reeder, F.; Chottard, J. C.; Dorsselaer, A. *Anal. Biochem.* **2000**, *284*, 77-86.

(23) Wu, Q.; Cheng, X.; Hofstadler, S. A.; Smith, R. D. *J. Mass Spectrom.* **1996**, *31*, 669-675.

(24) Favre, A.; Gonnet, F.; Tabet, J.-C. *Int. J. Mass Spectrom.* **1999**, *190*, 303-312.

(25) Hettich, R. L. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 941-949.

(26) Hettich, R. L. *Int. J. Mass Spectrom.* **2001**, *204*, 55-75.

(27) Wang, Y.; Taylor, J. S.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 550-556.

(28) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I. J. *Physical Chemistry of Nucleic Acids*; Harper and Row: New York, 1974.

(29) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.

(30) Wang, Z.; Wan, K. X.; Ramanathan, R.; Taylor, J. S.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 683-691.

(31) Wan, K. X.; Gross, J.; Hillenkamp, F.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 193-205.

(32) Favre, A.; Gonnet, F.; Tabet, J.-C. *Eur. J. Mass Spectrom.* **2000**, *6*, 389-396.

(33) Luo, H.; Lipton, M. S.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 195-199.

(34) Daneshfar, R.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 55-64.

(35) Hoaglund, C. S.; Liu, Y.; Ellington, A. D.; Pagel, M.; Clemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 9051-9052.

(36) Gidden, J.; Bushnell, J. E.; Bowers, M. T. *J. Am. Chem. Soc.* **2001**, *123*, 5610-5611.

(37) Gidden, J.; Bowers, M. T. *Eur. Phys. J. D* **2002**, *20*, 409-419.

(38) Gidden, J.; Bowers, M. T. *J. Phys. Chem. B* **2003**, *107*, 12829-12837.

(39) Gidden, J.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 161-170.

(40) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

(41) Keller, K. M.; Brodbelt, J. S. *Anal. Biochem.*. **2004**, *326*, 200-210.

(42) Bartlett, M. G.; McCloskey, J. A.; Manalili, S.; Griffey, R. H. *J. Mass Spectrom.* **1996**, *31*, 1277-1283.

(43) O'Hair, R. A. J.; McLuckey, S. A. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *162*, 183-202.

(44) Clauwaert, J.; Stockx, J. *Z. Naturforsch. B* **1968**, *23*, 25-30.

(1) Smith, D. L.; Deng, Y.; Zhang, Z. *J. Mass Spectrom.* **1997**, *32*, 135-146.

(2) Smith, R. D.; Bruce, J. E.; Wu, Q.; Lei, Q. P. *Chem. Soc. Rev.* **1997**, *26*, 191-202.

(3) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(4) Loo, J. A.; Loo, R. R. O. In *Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications*; Cole, R. B., Ed.; John Wiley & Sons: New York, 1997; pp 385-419.

(5) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(6) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, *120*, 9605-9613.

(7) Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. *Rapid Commun. in Mass Spectrom.* **2000**, *14*, 464-467.

(8) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(9) Gabelica, V.; De Pauw, E. *J. Mass Spectrom.* **2001**, *36*, 397-402.

(10) Gabelica, V.; De Pauw, E. *Int. J. Mass Spectrom.* **2002**, *219*, 151-159.

(11) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

(12) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.

(13) Goodlett, D. R.; Camp, D. G., II; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, *22*, 181-183.

(14) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328-1337.

(15) Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2489-2497.

(16) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(17) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.

(18) Carrasco, C.; Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E.; Garbay-Jaureguiberry, C.; Roques, B.; Wilson, W. D.; Chaires, J. B.; Waring, M. J.; Bailly, C. *ChemBioChem* **2002**, *3*, 1235-1241.

(19) David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.

(20) Guittat, L.; Alberti, P.; Rosu, F.; Van Miert, S.; Thetiot, E.; Pieters, L.; Gabelica, V.; De Pauw, E.; Ottaviani, A.; Riou, J.-F.; Mergny, J.-L. *Biochimie* **2003**, *85*, 535-547.

(21) Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

(22) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 10129-10133.

(23) Hofstadler, S. A.; Sannes-Lowery, K. A.; Crooke, S. T.; Ecker, D. J.; Sasmor, H.; Manalili, S.; Griffey, R. H. *Anal. Chem.* **1999**, *71*, 3436-3440.

(24) Griffey, R. H.; Sannes-Lowery, K. A.; Drader, J. J.; Mohan, V.; Swayze, E. E.; Hofstadler, S. A. *J. Am. Chem. Soc.* **2000**, *122*, 9933-9938.

(25) Sannes-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264-271.

(26) Sannes-Lowery, K. A.; Drader, J. J.; Griffey, R. H.; Hofstadler, S. A. *Proc. SPIE-Int. Soc. Opt. Eng.* **2001**, *4264*, 27-36.

(27) Hogan, J. M.; McLuckey, S. A. *J. Mass Spectrom.* **2003**, *38*, 245-256.

(28) Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal. Chem.* **2001**, *73*, 3274-3281.

(29) Newton, K. A.; Chrisman, P. A.; Reid, G. E.; Wells, J. M.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2001**, *212*, 359-376.

(30) Wells, J. M.; Stephenson, J. L.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2000**, *203*, A1-A9.

(31) McLuckey, S. A.; Vaidyanathan, G. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *162*, 1-16.

(32) Keller, K. M.; Brodbelt, J. S. *Anal. Biochem.* **2004**, *326*, 200-210.

(33) Kallenbach, N. R., Ed. *Chemistry and Physics of DNA-Ligand Interactions*; Adenine Press: Schenectady, NY, 1990.

(34) Propst, C. L.; Perun, T. J., Eds. *Nucleic Acid Targeted Drug Design*; M. Dekker: New York, 1992.

(35) Gupta, R.; Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 2472-2480.

(36) Zimmer, C.; Waehnert, U. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31-112.

(37) Uytterhoeven, K.; Sponer, J.; Van Meervelt, L. *Eur. J. Biochem.* **2002**, *269*, 2868-2877.

(38) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 5723-5727.

(39) Bailly, C.; Henichart, J. P.; Colson, P.; Houssier, C. *J. Mol. Recognit.* **1992**, *5*, 155-171.

(40) Bailly, C.; Colson, P.; Henichart, J. P.; Houssier, C. *Nucleic Acids Res.* **1993**, *21*, 3705-3709.

(41) Colson, P.; Houssier, C.; Bailly, C. *J. Biomol. Struct. Dyn.* **1995**, *13*, 351-366.

(42) Steinmetzer, K.; Reinert, K. E. *J. Biomol. Struct. Dyn.* **1998**, *15*, 779-791.

(43) Shen, L. L.; Baranowski, J.; Pernet, A. G. *Biochemistry* **1989**, *28*, 3879-3885.

(44) Shen, L. L.; Mitscher, L. A.; Sharma, P. N.; O'Donnell, T. J.; Chu, D. W.; Cooper, C. S.; Rosen, T.; Pernet, A. G. *Biochemistry* **1989**, *28*, 3886-3894.

(45) Marians, K. J.; Hiasa, H. *J. Biol. Chem.* **1997**, *272*, 9401-9409.

(46) Khodursky, A. B.; Cozzarelli, N. R. *J. Biol. Chem.* **1998**, *273*, 27668-27677.

(47) Kampranis, S. C.; Maxwell, A. *J. Biol. Chem.* **1998**, *273*.

(48) Kwok, Y.; Zeng, Q.; Hurley, L. H. *J. Biol. Chem.***1999**, *274*, 17226-17235.

(49) Son, G. S.; Yeo, J.-A.; Kim, M.-S.; Kim, S. K.; Holmen, A.; Akerman, B.; Norden, B. *J. Am. Chem. Soc.* **1998**, *120*, 6451-6457.

(50) Son, G. S.; Yeo, J.-A.; Kim, J.-M.; Kim, S. K.; Moon, H. R.; Nam, W. *Biophys. Chem.* **1998**, *74*, 225-236.

(51) Lee, E.; Yeo, J.-A.; Cho, C. B.; Lee, G. J.; Han, S. W.; Kim, S. K. *Eur. J. Biochem.* **2000**, *267*, 6018-6024.

(52) Lee, H. M.; Kim, J.-K.; Kim, S. K. *J. Biomol. Struct. Dyn.* **2002**, *19*, 1083-1091.

(1) Ellington, A. D. *Curr. Biol.* **1994**, *4*, 427-429.

(2) Conrad, R. C.; Bruck, F. M.; Bell, S.; Ellington, A. D. *RNA Prot. Interact.* **1998**, 285-325.

(3) Famulok, M.; Mayer, G.; Blind, M. *Acc. Chem. Res.* **2000**, *33*, 591-599.

(4) Hermann, T.; Patel, D. J. *Science* **2000**, *287*, 820-825.

(5) Hesselberth, J.; Robertson, M. P.; Jhaveri, S.; Ellington, A. D. *Rev. Mol. Biotechnol.* **2000**, *74*, 15-25.

(6) Osborne, S. E.; Ellington, A. D. *Chem. Rev.* **1997**, *97*, 349-370.

(7) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611-647.

(8) Burgstaller, P.; Jenne, A.; Blind, M. *Curr. Opin. Drug Disc. Dev.* **2002**, *5*, 690-700.

(9) Sullenger, B. A.; Gilboa, E. *Nature* **2002**, *418*, 252-258.

(10) Brodbelt, J. S. *Int. J. Mass Spectrom.* **2000**, *200*, 57-69.

(11) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.

(12) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

(13) Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.

(14) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

(15) Cassiday, L. A.; Lebruska, L. L.; Benson, L. M.; Naylor, S.; Owen, W. G.; Maher, L. J., 3rd *Anal. Biochem.* **2002**, *306*, 290-297.

(16) Cavanagh, J.; Benson, L. M.; Thompson, R.; Naylor, S. *Anal. Chem.* **2003**, *75*, 3281-3286.

(17) Wang, Y.; Rando, R. R. *Chem. Biol.* **1995**, *2*, 281-290.

(18) Jiang, L.; Patel, D. J. *Nat. Struct. Biol.* **1998**, *5*, 769-774.

(19) Huizenga, D. E.; Szostak, J. W. *Biochemistry* **1995**, *34*, 656-665.

(20) Lin, C. H.; Patel, D. J. *Chem. Biol.* **1997**, *4*, 817-832.

(21) Burgstaller, P.; Famulok, M. *Angew. Chem.* **1994**, *106*, 1163-1166.

(22) Fan, P.; Suri, A. K.; Fiala, R.; Live, D.; Patel, D. J. *J. Mol. Biol.* **1996**, *258*, 480-500.

(23) Schneider, C.; Suhnel, J. *Biopolymers* **1999**, *50*, 287-302.

(24) Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B. *Science* **1994**, *263*, 1425-1429.

(25) Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J.-P.; Pardi, A. *Nat. Struct. Biol.* **1997**, *4*, 644-649.

(26) Zimmermann, G. R.; Wick, C. L.; Shields, T. P.; Jenison, R. D.; Pardi, A. *RNA* **2000**, *6*, 659-667.

(27) In.

(28) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457.

(29) Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300-307.

(30) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98.

(31) Keller, K. M.; Oehlers, L.; Brodbelt, J. S. **2004**, manuscript in preparation.

Vita

Karin Mia Keller was born in Oshkosh, Wisconsin on May 16, 1968 to Katherine J. and Dr. Meredith E. Keller. She received a diploma from Naperville Central High School in Naperville, Illinois (1986) and a Bachelor's Degree in Chemistry from Trinity University in San Antonio, Texas (1990). She was employed in the Medicinal Chemistry Department at Texas Biotechnology Corporation in Houston, Texas from 1992 to 1999 and co-authored several publications and patents during that period. In 1999 she entered the Graduate School of The University of Texas at Austin and joined the research group of Prof. Jennifer S. Brodbelt in the Department of Chemistry and Biochemistry.

Permanent address: 6803 Poncha Pass, Austin, Texas 78749

This dissertation was typed by the author.